# On the estimation of the variability in the distribution tail

Laurent Gardes, Stéphane Girard

# On the estimation of the variability in the distribution tail

Laurent Gardes[1,*] and Stéphane Girard[2]

[1] Institut de Recherche Mathématique Avancée, UMR 7501 Université de Strasbourg et CNRS, 7 rue René Descartes, 67000 Strasbourg, France.
[2] Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France.
[*] gardes@unistra.fr (corresponding author)

**Abstract**

We propose a new measure of variability in the tail of a distribution by applying a Box-Cox transformation of parameter $p \geq 0$ to the tail-Gini functional. It is shown that the so-called Box-Cox Tail Gini Variability measure is a valid variability measure whose condition of existence may be as weak as necessary thanks to the tuning parameter $p$. The tail behaviour of the measure is investigated under a general extreme-value condition on the distribution tail. We then show how to estimate the Box-Cox Tail Gini Variability measure within the range of the data. These methods provide us with basic estimators that are then extrapolated using the extreme-value assumption to estimate the variability in the very far tails. The finite sample behavior of the estimators is illustrated both on simulated and real data.

**Keywords**: Gini functional; risk measure; variability measure; distribution tail; extreme-value theory.

## 1 Introduction

The assessment of extreme risk is a crucial question in actuarial science and finance [17, 28, 38]. The most popular tail-based risk measure is the value-at-risk. It is defined as the quantile $Q(\alpha)$ of a real-valued random variable $X$ at level $\alpha \in (0,1)$. Here, we adopt the convention that $Q(\alpha)$ is the lowest value exceeded by $X$ with probability not larger than $\alpha$. As a consequence, the value-at-risk only provides an information about the frequency of an extreme event but not on its actual magnitude. To overcome this limitation, several alternative risk measures have been introduced. A first direction of research consists in developing statistical analogues of quantiles but taking into account the whole distribution tail: Expectiles [6, 10, 12, 25, 32], $L^p$-quantiles [8, 9, 11] or $L^p$-medians [20]. The interpretation of these risk measures can rely on the remark that expectiles, and more generally $L^p$-quantiles, coincide with quantiles of a transformed distribution [27]. A second line of work is to average the values of $X$ exceeding the value-at-risk. This gives rise to the conditional tail expectation or, when $X$ is continuous, the so-called expected-shortfall also known as tail value-at-risk, or conditional value-at-risk, see [3, 36, 39]. To take into account the variability of $X$ above the value-at-risk $Q(\alpha)$, the tail-standard-deviation has been introduced in [18] as a linear combination of the conditional tail expectation and the standard-deviation measure. This definition although requires the second moment of the distribution of $X$ to be finite, which may not be the case in practice, see for instance [34, Chapter 4]. It has then be proposed in [19] to replace the standard deviation measure by the tail-Gini functional defined as $\text{TGini}_X(\alpha) = \mathbb{E}\left\{|X - X^*| \big| \min(X, X^*) > Q(\alpha)\right\}$, where $X^*$ is an independent copy of $X$. It is shown in particular that the tail-Gini functional is a valid variability measure as soon as the first moment of the distribution of $X$ is finite.

We propose an extension of the tail-Gini functional obtained by replacing $|X - X^*|$ by its Box-Cox transformation of parameter $p \geq 0$. The so-called Box-Cox Tail Gini Variability measure $G_X(\alpha; p)$ is still a valid variability measure whose condition of existence is weakened compared to $\text{TGini}_X(\alpha)$ when $0 \leq p < 1$. The tail behaviour of $G_X(\alpha; p)$ is investigated for levels $\alpha \to 0$ under a general extreme-value condition on the distribution tail of $X$. We then show how to estimate the Box-Cox Tail Gini Variability measure $G_X(\alpha_n; p)$ at a level $\alpha_n \to 0$ as the sample size $n \to \infty$. Two estimation methods are first introduced for intermediate levels $\alpha_n$ such that $n\alpha_n \to \infty$. These methods provide us with basic estimators that are then extrapolated at an extreme level $\alpha_n$ such as $n\alpha_n \to c \in [0, \infty)$, using an extreme-value assumption.

The paper is organised as follows. The Box-Cox Tail Gini Variability measure is defined in Section 2 and some elementary properties are established. Section 3 then focuses on the tail properties (as $\alpha \to 0$) at the population level. The estimation of the extreme Box-Cox Tail Gini Variability measure is discussed in Section 4. A simulation study of the finite-sample performance of the estimators is presented in Section 5, and two applications to reinsurance data are discussed in Section 6. Proofs and auxiliary results are postponed to Section 7.

## 2 Box-Cox Tail Gini Variability measure

Let $X$ be a real-valued random variable with associated cumulative distribution function $F$. The corresponding quantile function (also referred to as the value-at-risk) is defined for $\alpha \in (0, 1)$ by $Q(\alpha) := F^{\leftarrow}(1 - \alpha)$. Let us recall that, for all increasing function $f$, the generalized inverse $f^{\leftarrow}$ is defined by $f^{\leftarrow}(t) = \inf\{x \in \mathbb{R};\ f(x) \geq t\}$, with $t \in \mathbb{R}$. It is assumed throughout the paper that $F$ is differentiable. The Box-Cox transformation [7] of parameter $p \geq 0$ is the function given for all $s > 0$ by $K_p(s) := \int_1^s u^{p-1} du$. We also refer to [29] for an application of this transformation to the estimation of regression quantiles.

**Definition 1** *Let $X^*$ be an independent copy of $X$. The Box-Cox Tail Gini Variability measure of parameter $p$ ($BC_p$-TGV measure) and level $\alpha \in (0, 1)$ is defined by*

$$G_X(\alpha; p) := K_p^{\leftarrow} \left(\mathbb{E}\left[K_p(|X - X^*|) \mid \min(X, X^*) > Q(\alpha)\right]\right),$$

*provided that the expectation exists.*

The above defined measure can be rewritten as

$$G_X(\alpha; p) = K_p^{\leftarrow} \left\{ \int_{(0,1)^2} K_p\left(|Q(\alpha u) - Q(\alpha v)|\right) du dv \right\},$$

or, more specifically,

$$G_X(\alpha; p) = \left\{ \int_{(0,1)^2} |Q(\alpha u) - Q(\alpha v)|^p \, du dv \right\}^{1/p}, \quad \text{if } p > 0, \tag{1}$$

$$G_X(\alpha; 0) = \exp\left\{ \int_{(0,1)^2} \ln |Q(\alpha u) - Q(\alpha v)| \, du dv \right\}, \quad \text{otherwise.}$$

In the particular case where $p = 1$ and $\mathbb{E}(|X|) < \infty$, the BC$_1$-TGV measure reduces to the tail-Gini functional $\mathrm{TGini}_X(\alpha)$ introduced in [19]. When $p = 2$ and $\mathbb{E}(X^2) < \infty$, one has

$$G_X^2(\alpha; 2) = 2[\mathrm{CTM}_X(\alpha; 2) - \mathrm{CTM}_X^2(\alpha; 1)] \tag{2}$$

where $\mathrm{CTM}_X(\alpha; r) = \mathbb{E}[X^r \big| X > Q(\alpha)]$ is the conditional tail moment [14], which can be interpreted as a particular Wang measure [15]. In the case where $r = 1$, $\mathrm{CTM}_X(\alpha; 1)$ is also called the tail conditional expectation risk measure and coincides with the expected shortfall when $F$ is continuous. Let us recall that, unlike value-at-risk, expected shortfall is a coherent risk measure, see [2, 36]. If follows from (2) that $G_X^2(\alpha; 2) = 2\,\mathrm{CTV_X}(\alpha)$ where $\mathrm{CTV_X}(\alpha)$ is the conditional tail variance introduced in [40]. Equivalently, $G_X(\alpha; 2) = \sqrt{2}\,\mathrm{SD_X}(\alpha)$ where $\mathrm{SD_X}(\alpha)$ is the tail standard-deviation measure, considered in [18]. We shall show in the following that choosing $p < 1$ in the BC$_p$-TGV measure yields valid measures of variability with associated conditions of existence weaker than $\mathrm{SD_X}(\alpha)$ and $\mathrm{TGini}_X(\alpha)$. We start by giving sufficient conditions for the existence of the BC$_p$-TGV measure.

**Proposition 1** *If $p > 0$ with $\mathbb{E}(\max(X, 0)^p) < \infty$ or $p = 0$ with $\mathbb{E}(|\ln|X||) < \infty$ and the density of $X$ is bounded, then*

$$\mathbb{E}\left[|K_p(|X - X^*|)| \mid \min(X, X^*) > Q(\alpha)\right] < \infty.$$

As stated in [19], properties of variability measures may be quite different from those of risk measures [4, 31]. They also slightly differ from those of deviation measures [37].

**Proposition 2** *When it exists, the BC$_p$-TGV measure of level $\alpha \in (0, 1)$ satisfies the following properties:*

*(i) Law invariance: If $X \stackrel{d}{=} Y$ (i.e. $X$ and $Y$ have the same distribution) then $G_X(\alpha; p) = G_Y(\alpha; p)$.*

*(ii) Standardization: $G_\mu(\alpha; p) = 0$ for all $\mu \in \mathbb{R}$.*

*(iii) Location invariance: $G_{X+\mu}(\alpha; p) = G_X(\alpha; p)$ for all $\mu \in \mathbb{R}$.*

*(iv) Positive homogeneity: $G_{\lambda X}(\alpha; p) = \lambda G_X(\alpha; p)$ for all $\lambda > 0$.*

Proposition 2$(i, ii, iii)$ imply that the BC$_p$-TGV measure is a valid measure of variability according to [19, Definition 2.3]. However, the BC$_p$-TGV measure is not coherent since the sub-additivity property is not fulfilled. As a comparison, $\mathrm{CTV_X}(\alpha)$ and $\mathrm{SD_X}(\alpha)$ are both non-coherent variability measures, and $\mathrm{CTV_X}(\alpha)$ does not fulfill $(iv)$.

**Proposition 3** *The set of measures $\{\mathrm{BC}_p{-}\mathrm{TGV}, p \geq 0\}$ is ordered: $0 \leq p \leq q$ implies $G_X(\alpha; p) \leq G_X(\alpha; q)$, whenever the measures exist.*

# 3 Tail behavior of the BC$_p$-TGV measure

Let $U(x) := Q(1/x)$ be the tail quantile function. From now on, it is assumed that $U$ satisfies the first order condition

**(C1)** There exist a positive function $a$ and $\gamma \in \mathbb{R}$ such that for all $s > 0$,

$$\lim_{x\to\infty} \Delta(s, x) = 0, \quad \text{where } \Delta(s, x) := \frac{U(sx) - U(x)}{a(x)} - K_\gamma(s).$$

In other words, $U$ is an extended regularly varying function with index $\gamma$. Condition **(C1)** is equivalent to assuming that the distribution of $X$ is in the maximum domain of attraction of the extreme-value distribution with extreme-value index $\gamma \in \mathbb{R}$, see [23, Theorem 1.1.6]. The next result provides a first order equivalent of the BC$_p$-TGV measure for extreme levels $\alpha \to 0$.

**Proposition 4** *Assume* **(C1)** *holds and*

$$\lim_{x\to\infty} \sup_{s>1} \max\left( \left| \frac{\Delta(s,x)}{K_\gamma(s)} \right|, \left| s^{-\gamma} \frac{a(sx)}{a(x)} - 1 \right| \right) = 0. \tag{3}$$

*Let $p \geq 0$ such that $p\gamma < 1$ and assume that $G_X(\alpha;p)$ exists for all $\alpha$ in a neighbourhood of 0. Then,*

$$\lim_{\alpha\to 0} \frac{G_X(\alpha;p)}{a(\alpha^{-1})} = \theta(p;\gamma),$$

*with, for $p \geq 0$ and $p\gamma < 1$,*

$$\theta(p;\gamma) := K_p^{\leftarrow} \left\{ \int_{(0,1)^2} K_p \left( \left| K_\gamma(u^{-1}) - K_\gamma(v^{-1}) \right| \right) du dv \right\}.$$

Note that when $p > 0$, $p\gamma < 1$ is a sufficient condition for $\mathbb{E}(\max(X,0)^p) < \infty$ and thus for the existence of the BC$_p$-TGV measure, see Proposition 1. The first part of condition (3) can be interpreted as a uniform version of **(C1)** while the second part of (3) is a strengthened regular variation property for the auxiliary function $a(\cdot)$, see [23, Theorem 2.3.3] and the discussion following Proposition 5 below. It appears that the tail behavior of the measure is mainly driven by the function $a(\cdot)$. Two opposite cases appear: When the distribution is in Fréchet maximum domain of attraction, it is heavy-tailed with extreme-value index $\gamma > 0$, $a(x) \sim \gamma U(x)$ as $x \to \infty$ and thus $G_X(\alpha;p) \to \infty$ as $\alpha \to 0$. The asymptotic variability is infinite in the distribution (heavy) tail. Conversely, in the situation where the distribution is in Weibull maximum domain of attraction, it is short-tailed with extreme-value index $\gamma < 0$, $a(x) \sim -\gamma(x_F - U(x))$ as $x \to \infty$, where $x_F$ is the finite right endpoint of the distribution of $X$. Hence, $G_X(\alpha;p) \to 0$ as $\alpha \to 0$. Asymptotically, there is no variability in the distribution (short) tail. The intermediate situation is also possible. The function $a(\cdot)$ associated with the Exponential distribution (in Gumbel maximum domain of attraction, with extreme-value index $\gamma = 0$) is constant, leading to an asymptotically finite dispersion in the (light) tail. Several explicit expressions of $\theta(p;\gamma)$ can be obtained depending on the sign of the extreme-value index $\gamma$. Introducing for $x > 0$ and $y > 0$,

$$B(x,y) = \int_0^1 t^{y-1}(1-t)^{x-1} dt, \quad \Gamma(x) = \int_0^\infty t^{x-1} \exp(-t) dt, \quad \psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$$

the beta, gamma and digamma functions, one has for $p \geq 0$ and $p\gamma < 1$,

$$K_p(\theta(p;\gamma)) = \begin{cases} (2B(p+1,\gamma^{-1}-p)/[\gamma^{p+1}(2-p\gamma)] - 1)/p & \text{if } \gamma > 0, \\ (\Gamma(p+1) - 1)/p & \text{if } \gamma = 0, \\ (2B(p+1,-\gamma^{-1})/[(-\gamma)^{p+1}(2-p\gamma)] - 1)/p & \text{if } \gamma < 0, \end{cases}$$

where, when $p = 0$, the right-hand side should be read as its limit when $p$ tends to 0, that is to say as

$$\begin{cases} \gamma/2 - \ln(\gamma) + \psi(1) - \psi(\gamma^{-1}) & \text{if } \gamma > 0, \\ \psi(1) & \text{if } \gamma = 0, \\ \gamma/2 - \ln(-\gamma) + \psi(1) - \psi(1 + (-\gamma)^{-1}) & \text{if } \gamma < 0. \end{cases}$$

Note also that $\theta(1;\gamma) = 2/((1-\gamma)(2-\gamma))$ for all $\gamma < 1$ leading to an asymptotic equivalent of the tail-Gini functional as $\alpha \to 0$:

$$\text{TGini}_X(\alpha) \sim \frac{2}{(1-\gamma)(2-\gamma)} a(\alpha^{-1}).$$

The graph of $\gamma \mapsto \theta(p;\gamma)$ is depicted on the top panel of Figure 1 for $\gamma \in [-1,1]$, $p \in \{0, 1/2, 1, 3/2, 2\}$ and under the constraint $p\gamma < 1$. It appears that the function $\gamma \mapsto \theta(p;\gamma)$ is increasing with respect to $\gamma$ for all considered values of $p$. Besides, $\ln\theta(0;\cdot)$ is approximately linear on the considered interval $\gamma \in [-1,1]$ while $\ln\theta(p;\cdot)$ is convex when $p > 0$. The rate of convergence of $G_X(\alpha;p)/a(\alpha^{-1})$ to $\theta(p;\gamma)$ in Proposition 4 can be established by considering a strengthened yet classical version of condition **(C1)**.

**(C2)** There exist a function $A$, not changing sign eventually such that $A(x) \to 0$ as $x \to \infty$ and some $\rho < 0$ such that

$$\lim_{x\to\infty} \sup_{s\in\mathcal{K}} |R(s,x)| = 0$$

for all compact subset $\mathcal{K}$ of $(1,\infty)$ and where

$$\begin{aligned} R(s,x) &:= A^{-1}(x)\left(\frac{U(sx)-U(x)}{a(x)} - K_\gamma(s)\right) - H_{\gamma,\rho}(s) \text{ and} \\ H_{\gamma,\rho}(s) &:= \int_1^s v^{\gamma-1}\int_1^v u^{\rho-1}\,du\,dv. \end{aligned}$$

Condition **(C2)** is a classical second-order condition on extended regularly varying functions. Its interpretation can be found in [5, 23] along with examples of distributions. For instance, in the case where $\gamma > 0$, Pareto, Burr, Fréchet, Student, Fisher and Inverse-Gamma distributions all satisfy this condition, and more generally so does the Hall-Weiss class of models [26]. Let for $s > t > 1$,

$$h_{\gamma,\rho}(s,t) := \frac{H_{\gamma,\rho}(s) - H_{\gamma,\rho}(t)}{K_\gamma(s) - K_\gamma(t)} \text{ and } \overline{R}(s,t,x) = \frac{R(s,x)-R(t,x)}{K_\gamma(s)-K_\gamma(t)}. \tag{4}$$

Clearly, $h_{\gamma,\rho}$ and $\overline{R}(\cdot,\cdot,x)$ are symmetric *i.e.* for all $(s,t) \in (1,\infty)^2$, $h_{\gamma,\rho}(s,t) = h_{\gamma,\rho}(t,s)$ and $\overline{R}(s,t,x) = \overline{R}(t,s,x)$. Moreover, for all $s > t > 1$, condition **(C2)** entails that $\overline{R}(s,t,x) \to 0$ as $x \to \infty$. Using the representation

$$\frac{G_X(\alpha;p)}{a(\alpha^{-1})} = K_p^{\leftarrow}\left\{\int_{(0,1)^2} K_p\left(\left|K_\gamma(u^{-1}) - K_\gamma(v^{-1})\right| W_\gamma(u,v)\right) du\,dv\right\},$$

where

$$W_\gamma(u,v) := \frac{U(\alpha^{-1}u^{-1}) - U(\alpha^{-1}v^{-1})}{(K_\gamma(u^{-1}) - K_\gamma(v^{-1}))a(\alpha^{-1})},$$

it is easily checked that, under **(C2)**

$$\begin{aligned} W_\gamma(u,v) &= 1 + A(\alpha^{-1})\left[h_{\gamma,\rho}\left(u^{-1},v^{-1}\right) + \overline{R}\left(u^{-1},v^{-1},\alpha^{-1}\right)\right] \\ &\approx 1 + A(\alpha^{-1})h_{\gamma,\rho}\left(u^{-1},v^{-1}\right). \end{aligned}$$

In view of the first order approximations $K_p(x+u) \approx K_p(x) + ux^{p-1}$ and $K_p^{\leftarrow}(x+u) \approx K_p^{\leftarrow}(x)[1+$

$u/(1+px)]$ for $x>0$ and $u$ close to 0, one can expect that for $\alpha$ close to 0

$$\frac{G_X(\alpha;p)}{a(\alpha^{-1})} \approx \theta(p;\gamma)\left\{1+A(\alpha^{-1})\frac{\mathcal{B}(p;\gamma)}{1+pK_p(\theta(p;\gamma))}\right\} = \theta(p;\gamma)+A(\alpha^{-1})\frac{\mathcal{B}(p;\gamma)}{\theta^{p-1}(p;\gamma)},$$

with, for all $p\geq 0$ and $p\gamma<1$,

$$\mathcal{B}(p;\gamma) := \int_{(0,1)^2} \left|K_\gamma(u^{-1})-K_\gamma(v^{-1})\right|^p h_{\gamma,\rho}\left(u^{-1},v^{-1}\right) dudv. \tag{5}$$

Note that the integral in the right hand-side term is well-defined since $h_{\gamma,\rho}$ is bounded on $(1,\infty)^2$ from Lemma 2 in Section 7. These approximations are rigorously justified in the following result.

**Proposition 5** *Assume* **(C2)** *holds and*

$$\lim_{x\to\infty}\sup_{s>1}\frac{|R(s,x)|}{K_\gamma(s)}=0, \tag{6}$$

*Let $p\geq 0$ such that $p\gamma<1$ and assume that $G_X(\alpha;p)$ exists for all $\alpha$ in a neighbourhood of 0. Then, as $\alpha\to 0$,*

$$\frac{G_X(\alpha;p)}{a(\alpha^{-1})} = \theta(p;\gamma)+A(\alpha^{-1})\mathcal{B}(p;\gamma)\theta^{1-p}(p;\gamma)(1+o(1)), \tag{7}$$

*where $\mathcal{B}(p;\gamma)$ is defined in (5).*

It can be proved that conditions **(C2)** and (6) of Proposition 5 imply conditions **(C1)** and (3) of Proposition 4. As a consequence of **(C2)**, the function $a(\cdot)$ is regularly varying at infinity with index $\gamma$, *i.e.* $a(sx)/a(x)\to s^\gamma$ as $x\to\infty$ for all $s>0$, see for instance [23, Theorem 2.3.3]. Proposition 5 thus entails that the BC$_p$-TGV measure inherits the regular variation property.

**Proposition 6** *Suppose* **(C2)** *holds. Let $p\geq 0$ such that $p\gamma<1$ and assume that $G_X(\alpha;p)$ exists for all $\alpha$ in a neighbourhood of 0. Then, $G_X(\cdot,p)$ is regularly varying at zero with index $-\gamma$.*

# 4 Inference

In this section, our goal is to estimate $G_X(\alpha_n,p)$ for levels $\alpha_n$ tending to zero at any rate, including both cases of intermediate $n\alpha_n\to\infty$ and extreme sequences $n\alpha_n\to c<\infty$. In view of the links established in Section 2, these estimators will also provide us with estimators of $\mathrm{TGini}_X(\alpha_n)$ and $\mathrm{SD}_X(\alpha_n)$. To this end, let $X_1,\dots,X_n$ be independent copies of $X$ and denote by $X_{1,n}\leq\ldots\leq X_{n,n}$ the associated order statistics. The first scenario is investigated in Subsection 4.1 while Subsection 4.2 shows how to extrapolate from an intermediate estimator to an extreme one. In the following, we shall use the notation $k_n=\lfloor n\alpha_n\rfloor$ where $\lfloor\cdot\rfloor$ stands for the floor function.

## 4.1 Estimation of the BC$_p$-TGV measure at intermediate levels

Let us first consider the estimation of $G_X(\alpha_n,p)$ where $(\alpha_n)$ is an intermediate sequence *i.e.* $\alpha_n\to 0$ and $n\alpha_n\to\infty$. Since, for $p\geq 0$,

$$G_X(\alpha_n;p)=K_p^{\leftarrow}\left\{\frac{1}{\alpha_n^2}\mathbb{E}\left[K_p\left(|X-X^*|\right)\mathbb{I}_{\{\min(X,X^*)>Q(\alpha_n)\}}\right]\right\},$$

a direct estimator of $G_X(\alpha_n, p)$ is obtained by considering the empirical counterpart of the mathematical expectation:

$$\widehat{G}_{X,n}(\alpha_n; p) := K_p^{\leftarrow}\left\{\frac{2}{k_n(k_n-1)}\sum_{i=1}^{k_n}\sum_{j=i+1}^{k_n} K_p\left(X_{n-i+1,n} - X_{n-j+1,n}\right)\right\}. \tag{8}$$

Note that, for all $p \geq 0$, the direct estimator satisfies properties $(ii)$, $(iii)$ and $(iv)$ of Proposition 2. The law invariance property $(i)$ is slightly modified as follows: $X \stackrel{d}{=} Y$ implies $\widehat{G}_{X,n}(\alpha_n; p) \stackrel{d}{=} \widehat{G}_{Y,n}(\alpha_n; p)$. The asymptotic normality of (8) is established in the next result. The proof consists in remarking that (8) can be interpreted as a triangular array of $U$-statistics: First, combining condition **(C2)** with Rényi's representation, the spacings between upper order statistics are approximated by some functions of exponential random variables. Second, the resulting $U$-statistics are controlled by [24, Theorem 7.1].

**Theorem 1** *Assume that* **(C2)** *and (6) hold with $p\gamma < 1/2$ and $p \geq 0$. Suppose that $G_X(\alpha; p)$ exists for all $\alpha$ in a neighbourhood of 0. Let $(\alpha_n)$ be an intermediate sequence and $k_n = \lfloor n\alpha_n \rfloor$ such that $k_n^{1/2} A(n/k_n) \to \lambda \in \mathbb{R}$. Then,*

$$k_n^{1/2}\left(\frac{\widehat{G}_{X,n}(\alpha_n; p)}{G_X(\alpha_n; p)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, \gamma^2 + 4\mathcal{V}(p;\gamma)\right),$$

*where, for $p \geq 0$,*

$$\mathcal{V}(p;\gamma) = \frac{1}{\theta^{2p}(p;\gamma)}\left\{\int_0^1\left[\int_0^1 K_p\left(\left|K_\gamma(u^{-1}) - K_\gamma(v^{-1})\right|\right) du\right]^2 dv - K_p^2(\theta(p;\gamma))\right\}.$$

Interestingly, the estimator $\widehat{G}_{X,n}(\alpha_n; p)$ is asymptotically unbiased even in the situation where $\lambda \neq 0$. The rate of convergence $k_n^{1/2}$ is driven by the effective number of observations used in (8). Note that condition $p\gamma < 1/2$ ensures the quantity $\mathcal{V}(p;\gamma)$ to be well-defined. The graph of the asymptotic variance $\gamma^2 + 4\mathcal{V}(p;\gamma)$ is depicted on the bottom panel of Figure 1 for $\gamma \in [-1, 1]$, $p \in \{0, 1/2, 1, 3/2, 2\}$ and under the constraint $p\gamma < 1/2$. It appears that the asymptotic variance is approximately independent of $p$ when $\gamma < -1/4$. In the case where $\gamma > -1/4$, the smallest asymptotic variance is obtained for $p = 0$.

Second, based on Proposition 4, a first idea to estimate the BC$_p$-TGV measure is to use the plug-in estimator $\theta(p, \hat{\gamma}_n)\hat{a}_n(\alpha_n^{-1})$ where $\hat{\gamma}_n$ and $\hat{a}_n(\alpha_n^{-1})$ are some consistent estimators of $\gamma$ and $a(\alpha_n^{-1})$. However, this estimator is well-defined only when $p\hat{\gamma}_n < 1$. In practice, even though $\hat{\gamma}_n$ converges in probability to $\gamma$ with $p\gamma < 1$, one can face the situation where $p\hat{\gamma}_n \geq 1$ when $p > 0$. To overcome this problem, we propose to use the statistics

$$\widetilde{G}_{X,n}(\alpha_n; p) := \theta(p, \hat{\gamma}_n^*)\hat{a}_n(\alpha_n^{-1}), \tag{9}$$

where $\hat{a}_n(\alpha_n^{-1})$ is some convenient estimator of $a(\alpha_n^{-1})$ and, for a consistent estimator $\hat{\gamma}_n$ of $\gamma$,

$$\hat{\gamma}_n^* := \min\left(\hat{\gamma}_n, \frac{2}{p} - \hat{\gamma}_n\right) \text{ if } p > 0 \text{ and } \hat{\gamma}_n^* := \hat{\gamma}_n \text{ if } p = 0. \tag{10}$$

Unlike the direct estimator defined in (8), the indirect estimator (9) does not necessarily satisfy the properties of variability measures listed in Proposition 2. The asymptotic normality of the indirect estimator is established for $p \geq 0$ in the following result. Since (9) is essentially a plug-in estimator, the proof of its asymptotic properties relies on a delta-method technique adapted to the extreme-value context.

**Theorem 2** *Assume that* **(C2)** *and (6) hold with $p\gamma < 1$ and $p \geq 0$. Suppose that $G_X(\alpha; p)$ exists for all $\alpha$ in a neighbourhood of 0. Let $(\alpha_n)$ be an intermediate sequence and $k_n = \lfloor n\alpha_n \rfloor$ such that $k_n^{1/2} A(n/k_n) \to \lambda \in \mathbb{R}$. If there exists a random pair $(\Lambda_1, \Lambda_2)$ such that*

$$k_n^{1/2} \left( \hat{\gamma}_n - \gamma, \frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})} - 1 \right) \xrightarrow{d} (\Lambda_1, \Lambda_2), \tag{11}$$

*then,*

$$k_n^{1/2} \left( \frac{\widetilde{G}_{X,n}(\alpha_n; p)}{G_X(\alpha_n; p)} - 1 \right) \xrightarrow{d} \frac{\dot{\theta}(p; \gamma)}{\theta(p; \gamma)} \Lambda_1 + \Lambda_2 - \lambda \frac{\mathcal{B}(p; \gamma)}{\theta^p(p; \gamma)},$$

*where $\dot{\theta}(p; \cdot)$ is the first derivative of the function $\gamma \mapsto \theta(p; \gamma)$.*

Let us note that $\hat{\gamma}_n^*$ inherits its asymptotic properties from $\hat{\gamma}_n$: Under (11), it is shown in the proof of Theorem 2 that

$$k_n^{1/2} \left( \hat{\gamma}_n^* - \gamma, \frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})} - 1 \right) \xrightarrow{d} (\Lambda_1, \Lambda_2).$$

In contrast to Theorem 1, the indirect estimator is asymptotically biased in the case where $\lambda \neq 0$. Three sources of bias appear in the asymptotic distribution: The asymptotic biases of $\hat{a}_n(\alpha_n^{-1})$ and $\hat{\gamma}_n$ as well as the remainder term in (7). However, compared to the direct estimator, the condition on $(p, \gamma)$ is weakened to $p\gamma < 1$. As an example, a popular estimator of $\gamma \in \mathbb{R}$ is the moment estimator [13] defined by

$$\hat{\gamma}_n^{(M)} := M_n^{(1)} + \hat{\gamma}_n^{(M,-)} \text{ with } \hat{\gamma}_n^{(M,-)} := 1 - \frac{1}{2} \left( 1 - \frac{(M_n^{(1)})^2}{M_n^{(2)}} \right)^{-1},$$

and where for $j \in \{1, 2\}$,

$$M_n^{(j)} := \frac{1}{k_n} \sum_{i=1}^{k_n - 1} \left[ \ln \left( \frac{X_{n-i+1,n}}{X_{n-k_n,n}} \right) \right]^j.$$

The corresponding estimator of $a(\alpha_n^{-1})$ (see for instance [23, Eq. (4.2.4)]) is

$$\hat{a}_n^{(M)} := X_{n-k_n,n} M_n^{(1)} \left( 1 - \hat{\gamma}_n^{(M,-)} \right).$$

Other possible estimators include the maximum likelihood estimator (if $\gamma > -1/2$, see for instance [23, Section 3.4]) or the probability-weighted moment estimator (if $\gamma < 1$, see for instance [23, Section 3.6.1]). The moment estimator of $G_X(\alpha_n; p)$ obtained by plugging $\hat{\gamma}_n^{(M)}$ and $\hat{a}_n^{(M)}$ in (9) is denoted by $\widetilde{G}_{X,n}^{(M)}(\alpha_n; p)$. It appears that the moment estimator of the BC$_p$-TGV measure satisfies properties (*ii*) and (*iv*) of Proposition 2. The law invariance (*i*) is also obtained up to the adaptation of the definition already discussed. On the contrary, the location invariance (*iii*) does not hold since $\hat{\gamma}_n^{(M)}$ and $\hat{a}_n^{(M)}$ are not location invariant. The asymptotic normality is a direct consequence of Theorem 2 and [23, Corollary 4.2.2].

**Corollary 1** *Assume that* **(C2)** *and (6) hold with $p\gamma < 1$, $p \geq 0$ and $\gamma \neq \rho$. Suppose that $G_X(\alpha; p)$ exists for all $\alpha$ in a neighbourhood of 0. Let $(\alpha_n)$ be an intermediate sequence and $k_n = \lfloor n\alpha_n \rfloor$ such* *that* $k_n^{1/2}A(n/k_n) \to \lambda \in \mathbb{R}$*. Then,*

$$k_n^{1/2}\left(\frac{\widetilde{G}^{(M)}_{X,n}(\alpha_n;p)}{G_X(\alpha_n;p)} - 1\right) \xrightarrow{d} \mathcal{N}\left(\lambda\mu^{(M)}(\gamma,\rho), \sigma^{(M)}(\gamma)^2\right),$$

*with*

$$\begin{aligned}
\mu^{(M)}(\gamma,\rho) &:= \frac{\dot{\theta}(p;\gamma)}{\theta(p;\gamma)}\mu_1(\gamma,\rho) + \mu_2(\gamma,\rho) - \frac{\mathcal{B}(p;\gamma)}{\theta^p(p;\gamma)}, \\
\sigma^{(M)}(\gamma)^2 &:= \left(\frac{\dot{\theta}(p;\gamma)}{\theta(p;\gamma)}\right)^2 v_1(\gamma) + 2\frac{\dot{\theta}(p;\gamma)}{\theta(p;\gamma)}v_{1,2}(\gamma) + v_2(\gamma),
\end{aligned}$$

*and where we have defined*

$$\begin{aligned}
\mu_1(\gamma,\rho) &:= \begin{cases} \frac{(1-\gamma)(1-2\gamma)}{(1-\gamma-\rho)(1-2\gamma-\rho)} & \text{if } \gamma < \rho < 0, \\ \frac{\gamma-\gamma\rho+\rho}{(\gamma+\rho)(1-\rho)^2} & \text{if } (0 < \gamma < -\rho \text{ and } \ell = 0) \text{ or } \gamma > -\rho > 0, \\ 0 & \text{otherwise}, \end{cases} \\
\mu_2(\gamma,\rho) &:= \begin{cases} \frac{\rho}{(1-\gamma-\rho)(1-2\gamma-\rho)} & \text{if } \gamma < \rho < 0, \\ \frac{\rho^2}{(\gamma+\rho)(1-\rho)^2} & \text{if } (0 < \gamma < -\rho \text{ and } \ell = 0) \text{ or } \gamma > -\rho > 0, \\ 0 & \text{otherwise}, \end{cases}
\end{aligned}$$

*with, for* $\gamma > 0$*,* $\ell := \lim_{x\to\infty} U(x) - a(x)/\gamma$ *and*

$$\begin{aligned}
v_1(\gamma) &:= \begin{cases} 1+\gamma^2 & \text{if } \gamma \geq 0, \\ \frac{(1-\gamma)^2(1-2\gamma)(1-\gamma+6\gamma^2)}{(1-3\gamma)(1-4\gamma)} & \text{if } \gamma < 0, \end{cases} \\
v_2(\gamma) &:= \begin{cases} 2+\gamma^2 & \text{if } \gamma \geq 0, \\ \frac{2-16\gamma+51\gamma^2-69\gamma^3+50\gamma^4-24\gamma^5}{(1-2\gamma)(1-3\gamma)(1-4\gamma)} & \text{if } \gamma < 0, \end{cases} \\
v_{1,2}(\gamma) &:= \begin{cases} \gamma - 1 & \text{if } \gamma \geq 0, \\ \frac{(1-\gamma)^2(-1+4\gamma-12\gamma^2)}{(1-3\gamma)(1-4\gamma)} & \text{if } \gamma < 0. \end{cases}
\end{aligned}$$

## 4.2 Estimation of the BC$_p$-TGV measure at extreme levels

We are now interested in the estimation of $G_X(\beta_n;p)$ for an extreme level $\beta_n$ satisfying $n\beta_n \to c < \infty$. The basic idea is to extrapolate the estimate of the BC$_p$-TGV measure at an intermediate level $\alpha_n$ such that $n\alpha_n \to \infty$, to the extreme level $\beta_n$. This is achieved by adapting Weissman's device [41] for estimating an extreme quantile to the BC$_p$-TGV framework. The regular variation property stated in Proposition 6 suggests that, under suitable conditions on $(\alpha_n, \beta_n)$, one may write

$$\frac{G_{X,n}(\beta_n;p)}{G_{X,n}(\alpha_n;p)} \approx \left(\frac{\beta_n}{\alpha_n}\right)^{-\gamma}.$$

This approximation leads to the following extrapolated estimator:

$$\overline{G}^{\star}_{X,n}(\beta_n;p) := \left(\frac{\beta_n}{\alpha_n}\right)^{-\hat{\gamma}_n}\overline{G}_{X,n}(\alpha_n;p), \tag{12}$$

where $\hat{\gamma}_n$ and $\overline{G}_{X,n}(\alpha_n; p)$ are some estimators of $\gamma$ and $G_X(\alpha_n; p)$. The asymptotic normality of the extrapolated estimator is given in the following result. The proof follows the same lines as the proof of the asymptotic normality of extreme quantile estimators, see for instance [23, Theorem 4.3.8] or [20, Theorem 2].

**Theorem 3** *Assume that* **(C2)** *and (6) hold with $p\gamma < 1$ and $p \geq 0$. Suppose that $G_X(\alpha; p)$ exists for all $\alpha$ in a neighbourhood of 0. Let $(\alpha_n)$ be an intermediate sequence and $k_n = \lfloor n\alpha_n \rfloor$ such that $k_n^{1/2} A(n/k_n) \to \lambda \in \mathbb{R}$. If there exists a random pair $(\Lambda_1, \Lambda_2)$ such that*

$$k_n^{1/2} \left( \hat{\gamma}_n - \gamma, \frac{\overline{G}_{X,n}(\alpha_n; p)}{G_X(\alpha_n; p)} - 1 \right) \xrightarrow{d} (\Lambda_1, \Lambda_2), \tag{13}$$

*then, for all sequence $(\beta_n)$ such that $\beta_n/\alpha_n \to 0$ and $k_n^{-1/2} \ln(\alpha_n/\beta_n) \to 0$,*

$$\frac{k_n^{1/2}}{\ln(\alpha_n/\beta_n)} \left( \frac{\overline{G}^*_{X,n}(\beta_n; p)}{G_X(\beta_n; p)} - 1 \right) \xrightarrow{d} \Lambda_1.$$

It appears that the limiting distribution of $\frac{\overline{G}^*_{X,n}(\beta_n;p)}{G_X(\beta_n;p)} - 1$ is the same as $\hat{\gamma}_n - \gamma$ with a different scaling. One can choose for $\overline{G}_{X,n}(\alpha_n; p)$ in (12) either the direct estimator $\widehat{G}_{X,n}(\alpha_n; p)$ or the indirect one $\widetilde{G}_{X,n}(\alpha_n; p)$. The corresponding extrapolated estimators of $G_X(\beta_n; p)$ are denoted by $\widehat{G}^*_{X,n}(\beta_n; p)$ and $\widetilde{G}^*_{X,n}(\beta_n; p)$ in the sequel. Let us note that the extrapolated indirect estimator can be re-interpreted as

$$\widetilde{G}^*_{X,n}(\beta_n; p) = \theta(p; \hat{\gamma}^*_n) \hat{a}^*_n(\beta_n^{-1}),$$

where $\hat{\gamma}^*_n$ is defined in (10) and

$$\hat{a}^*_n(\beta_n^{-1}) := \left( \frac{\beta_n}{\alpha_n} \right)^{-\hat{\gamma}_n} \hat{a}_n(\alpha_n^{-1})$$

is an extrapolated estimator of $a(\beta_n^{-1})$. Finally, considering the moment estimator $\hat{\gamma}_n^{(M)}$ in (12), the convergence in distribution of Theorem 3 holds with $\Lambda_1 = \mathcal{N}(\lambda\mu_1(\gamma, \rho), v_1(\gamma))$.

# 5 Illustration on simulated data

**Simulated model** The random variable $X$ of interest is such that for all $x \in \mathcal{X}$,

$$F(x) = 1 - \left[ K_\kappa^{\leftarrow}(x^c) \right]^{-1/c}, \tag{14}$$

where $\kappa \in \mathbb{R}$, $c > 0$ and $\mathcal{X} := \{x \geq 0;\ 1 + \kappa x^c > 0\}$. The corresponding tail quantile function defined for all $x \geq 1$ by $U(x) = [K_\kappa(x^c)]^{1/c}$ satisfies **(C1)** with $\gamma = \max(\kappa, 0) + c\min(\kappa, 0)$ and $a(x) = |\kappa|^{1-1/c} x^\gamma (1 - (1 - c^{-1}) x^{-c\kappa^2/|\kappa|})^{-1}$. Three main cases appear. If $\kappa > 0$, then $F$ is the cumulative distribution function of a Burr distribution in Fréchet maximum domain of attraction and $\gamma > 0$. If $\kappa = 0$, then $F$ is the cumulative distribution function of a Weibull distribution in Gumbel maximum domain of attraction and $\gamma = 0$. Finally, if $\kappa < 0$, then $X$ has a finite right endpoint, the associated distribution is in Weibull maximum domain of attraction and $\gamma < 0$. Moreover, if $c \neq 1$ and $\kappa \neq 0$, then conditions **(C2)** and (6) hold with $\rho = -c\kappa^2/|\kappa| < 0$ and $A(x) = \rho(1 - c^{-1}) x^\rho$. If $c \neq 1$ and $\kappa = 0$, then $A(x) = c(1 - c)^{-1} \ln(x)$ and $\rho = 0$, condition **(C2)** does not hold. If $c = 1$ and $\kappa = 0$, then one can set $A(x) = 0$ and $\rho = -\infty$ to fulfill conditions **(C2)** and (6).

To assess the finite sample performance of the estimators, $N = 500$ independent replications of a $n$-sample $X_1, \ldots, X_n$ drawn from (14) are generated with $n \in \{500, 5000\}$.

**Intermediate level** Here, the level of the BC$_p$-TGV measure is fixed to $\alpha_n = 3n^{-1/2}$. The values of the corresponding BC$_p$-TGV measure are represented in Figure 2 for $p \in \{0, 1/2, 1, 3/2, 2\}$ and $n = 5000$ as a function of $\kappa \in [-3/4, 1]$. Two values of $c$ are considered: $c = 1/2$ and $c = 2$. We compute $N$ independent realizations of the direct and indirect estimators of $G_X(\alpha_n; p)$ given by (8) and (9) for $p \in \{0, 1, 2\}$ and $c \in \{1/2, 2\}$. In the case of the indirect estimator, the moment estimators of $\gamma$ and $a(\alpha_n^{-1})$ are used. The median and the empirical quantiles of level 5% and 95% of the $N$ realizations are represented as functions of $\kappa \in [-3/4, 1]$ in Figure 3 ($n = 500$) and Figure 4 ($n = 5000$). It appears that, in all considered cases, the direct estimator provides better results than the indirect one, both in terms of bias and variance.

**Extreme level** We are now interested in the finite sample performance of the extrapolated estimator (12) of $G_X(\beta_n, p)$ with $\beta_n = 1/n$ and $\alpha_n = k_n/n$. As an estimator of $\gamma$, we use the moment estimator $\hat{\gamma}_n^{(M)} = \hat{\gamma}_n^{(M)}(k_n)$ and $\overline{G}_{X,n}(k_n/n; p)$ is either the direct or the indirect estimator of $G_X(k_n/n; p)$. For all $p \in \{0, 1, 2\}$, the tuning parameter $k_n$ is taken as

$$\hat{k}_{\mathrm{opt};p} := \underset{k \in \{16, \ldots, \lfloor n/4 \rfloor\}}{\arg\min} \ln^2 \left( \frac{\widehat{G}^*_{X,n}\,(k/(4n); p)}{\widehat{G}_{X,n}\,(k/(4n); p)} \right).$$

The idea motivating this choice is that, for a well-chosen value of $k_n$, the direct estimator and the extrapolated direct estimator should provide a similar estimation of the BC$_p$-TGV measure $G_X(k/(4n); p)$. This procedure is also used for instance in the framework of other risk measures based on $L^p-$optimization [20]. Simulation settings are the same as in the intermediate case. The results displayed in Figure 5 ($n = 500$) and Figure 6 ($n = 5000$) are satisfying. It appears that both direct and indirect extrapolated estimators yield similar results. This empirical observation can be explained by Theorem 3: The behavior of the extrapolated estimator is mainly driven by the estimator of the extreme-value index.

# 6 Real data examples

## 6.1 Norwegian fire losses data set

The data set consists in $n = 9181$ fire losses over the period 1972 to 1992. These data are available for instance in the **R** package `CASdatasets` as `data(norfire)`. The amount of losses are corrected for inflation using the Norwegian consumer price index. For each year $j \in \{1972, \ldots, 1992\}$, we denote by $X_1^{(j)}, \ldots, X_{n_j}^{(j)}$ the observations of the $n_j$ fire losses. These observations are assumed to be independent and generated from a parent random variable $X^{(j)}$. The data associated with year 1976 have been studied in details in [5, Example 1.2].

Here, we are interested in the comparison of the tail variabilities of the fire losses over the whole period 1972 to 1992. With this goal in mind, we estimate the BC$_p$-TGV measures $G_{X^{(j)}}(\beta; p)$ for given values of $\beta \in (0, 1)$ and $p > 0$. The year 1972 corresponds to the smallest sample with $\min(n_{1972}, \ldots, n_{1992}) = 97$ observations while 1988 provides the largest sample with 827 observations. The level $\beta = 1/100$ is selected which is extreme at least for the year 1972. We thus focus on the extrapolated versions of the direct and indirect estimators.

**Validity of condition (C1)** We first compute, for each year $j \in \{1972, \dots, 1992\}$, the moment estimator $\hat{\gamma}_n^{(j)}$ of the extreme-value indices $\gamma^{(j)}$, $j = 1972, \dots, 1992$. The number of observations used in the estimation procedure is fixed to $k_n^{(j)} = \lfloor n_j/6 \rfloor$. The smallest value of the moment estimator is obtained in 1980 ($\hat{\gamma}_n^{(1980)} \approx 0.256$) and the largest one in 1985 ($\hat{\gamma}_n^{(1985)} \approx 0.885$). It thus seems reasonable to assume that for each year, the fire losses distribution satisfies condition **(C1)** with a positive index. This assumption can be graphically checked on the QQ-plots

$$\left\{ \left( \ln k_n^{(j)} - \ln i, \; \ln X_{n_j - i + 1, n_j}^{(j)} - \ln X_{n_j - k_n^{(j)}, n_j}^{(j)} \right), i = 1, \dots, k_n^{(j)} \right\}.$$

Indeed, under **(C1)**, these plots must be approximately linear with positive slope $\gamma^{(j)}$. As an illustration, the QQ-plots associated with years 1980 and 1985 are displayed in left panel of Figure 7. They confirm the adequacy of **(C1)** to the data set.

**Choice of $p$** The value of $p$ must be chosen in order to satisfy the condition $p < 1/\tilde{\gamma}$ where $\tilde{\gamma} := \max\{\gamma^{(1972)}, \dots, \gamma^{(1992)}\}$. To select such a value, let

$$\bar{\gamma}_n^{(j)} := \hat{\gamma}_n^{(j)} + u_{0.99} \sqrt{v_1(\hat{\gamma}_n^{(j)})/k_n^{(j)}},$$

where $u_\alpha$ is the quantile of level $\alpha$ of the standard Gaussian distribution and $v_1(\cdot)$ is defined in Corollary 1. Based on the asymptotic normality of $\hat{\gamma}_n^{(j)}$, see for instance [13], the probability that $\bar{\gamma}_n^{(j)}$ is larger than the true extreme-value index $\gamma^{(j)}$ is, for $n_j$ large enough, approximately 0.99. Letting

$$\hat{p} := 1 \Big/ \max \left\{ \bar{\gamma}_n^{(1972)}, \dots, \bar{\gamma}_n^{(1992)} \right\} \approx 0.694,$$

and assuming that the sub-samples $\{X_1^{(j)}, \dots, X_{n_j}^{(j)}\}$, $j = 1972, \dots, 1992$ are independent,

$$\mathbb{P}\left(\tilde{\gamma} < 1/\hat{p}\right) = 1 - \prod_{j=1972}^{1992} \mathbb{P}(\bar{\gamma}_n^{(j)} > \tilde{\gamma}) \geq 1 - \prod_{j=1972}^{1992} \mathbb{P}(\bar{\gamma}_n^{(j)} > \gamma^{(j)}) \approx 1 - (1 - 0.99)^{21}.$$

As a consequence, the condition $\hat{p} < 1/\tilde{\gamma}$ is satisfied with a probability larger than $1 - (1 - 0.99)^{21} \approx 1$. We thus propose to estimate the BC$_p$-TGV measure with $p = \hat{p} \approx 0.694$.

Since $\hat{p} < 1$, it appears that for this data set, the tail variability cannot be measured neither by the tail Gini functional [19] nor by the tail standard-deviation [18]. At the opposite, it is possible to compute the direct and indirect extrapolated estimators of the BC$_p$-TGV measure for each year. The value of the intermediate sequence is selected by the procedure presented in Section 5 and the results are depicted in the left panel of Figure 8. As expected, both estimators yield similar results. Note also that the estimated values of the BC$_p$-TGV measure are in adequation with the visual variability within the data set.

## 6.2 Danish fire losses data set

This second data set consists of fire losses collected at Copenhagen Reinsurance over the period 1980 to 1990. This data set (available in the **R** package `CASdatasets`) has been widely studied in the actuarial literature (see for instance [1, 21, 22, 30, 33] among others). Here, $n = 2167$ losses (in millions of Danish Krone) were recorded and adjusted for inflation. We adopt the notations introduced in the previous section. The smallest sample size is $n_{1983} = 153$ while the largest is $n_{1986} = 238$. We consider the estimation of the BC$_p$-TGV measure for a level $\beta = 1/150$ which is extreme for the year 1983. For $j \in \{1980, \ldots, 1990\}$, taking $k_n^{(j)} = \lfloor n_j/6 \rfloor$ order statistics, the moment estimators of $\gamma^{(j)}$ are such that $0 < \hat{\gamma}_n^{(1983)} \approx 0.299 \leq \hat{\gamma}_n^{(j)} \leq \gamma_n^{(1980)} \approx 0.86$. As it was the case for the Norwegian fire losses data set, one thus may be confident in the validity of **(C1)** with a positive index; see also the QQ-plots in the right panel of Figure 7. Note that the large value of the estimated extreme value index for the year 1980 is essentially due to the presence of a very large fire loss. We first choose the value of the power $p$ of the BC$_p$-TGV measure. Using the procedure introduced in the previous section, the value $\hat{p} \approx 0.724 < 1$ is selected. The two extrapolated estimators (direct and indirect) are next depicted in the right panel of Figure 8. It is worth noting that the estimated BC$_p$-TGV measure is large in 1980. This is due to the important fire loss recorded this year. The effect of such extreme data points could be reduced by considering trimmed or winsorised estimators of BC$_p$-TGV measure, see [16] for the case of extreme Wang distortion risk measures.

# 7 Proofs

## 7.1 Preliminary results

Our first four results are of analytical nature: They provide upper and lower bounds on the function $h_{\gamma,\rho}$ as well as a uniform convergence result for $\overline{R}(s,t,\cdot)$ both defined in (4).

**Lemma 1** *For all $(x, y) \in (1, \infty)^2$ and all $c > 0$,*

$$\min(1-c, 0) \leq \frac{x^{1-c} - y^{1-c}}{x - y} \leq \max(1-c, 0), \quad 0 \leq \frac{x^{-(1+c)} - y^{-(1+c)}}{x^{-1} - y^{-1}} \leq 1 + c,$$
$$0 \leq \frac{\ln(x/y)}{x-y} \leq 1, \ \textit{and} \qquad -c \leq \frac{x^{-c} - y^{-c}}{\ln(x/y)} \leq 0.$$

**Proof** − We only present the study of the function

$$g_c(x, y) := \frac{x^{1-c} - y^{1-c}}{x - y}.$$

The three other functions can be studied in a similar way. First, since $g_c(x, y) = g_c(y, x)$ one can assume in what follows that $x > y$. Denoting by $\dot{g}_c$ the partial derivative of $g_c$ with respect to its first argument, one has, for $x > y > 1$, that

$$(x-y)^2 \dot{g}_c(x, y) = (1-c)x^{-c}(x-y) - x^{1-c} + x^{1-c}\left(1 - \frac{x-y}{x}\right)^{1-c}.$$

Let us first focus on the case $c < 1$. In this situation, $g_c(x, y) > 0$ for all $x > y > 1$. Since for all $t \in [0, 1]$, $(1-t)^{1-c} \leq 1 - (1-c)t$, it follows that

$$(x-y)^2 \dot{g}_c(x, y) \leq (1-c)x^{-c}(x-y) - x^{1-c} + x^{1-c}\left(1 - (1-c)\frac{x-y}{x}\right) = 0.$$

Hence, the maximum of $g_c$ is reached when $x = y$. Since

$$\lim_{x \downarrow y} g_c(x, y) = (1-c)y^{-c},$$

is a decreasing function of $y$, the maximum is reached for $x = y = 1$ and is equal to $(1-c)$.
Now, if $c > 1$, then $g_c(x,y) < 0$. Furthermore, using the inequality $(1-t)^{1-c} \geq 1-(1-c)t$ that holds for all $t \in [0,1]$, one can remark that $\dot{g}_c(x,y) > 0$ for all $x > y$. Hence, the minimum of $g_c$ is reached for $x = y = 1$ and is equal to $(1-c)$. Collecting the previous two conclusions leads to the bounds provided in Lemma 1. ■

**Lemma 2** *The function $h_{\gamma,\rho}$ defined in (4) is bounded: $0 \leq h_{\gamma,\rho}(x,y) \leq -1/\rho$, for all $(x,y) \in (1,\infty)^2$, $\gamma \in \mathbb{R}$ and $\rho < 0$.*

**Proof** – Let us consider separately the three following situations.
$(i)$ When $\gamma \neq 0$ and $\gamma \neq -\rho$,

$$H_{\gamma,\rho}(x) = \frac{1}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}\right) \text{ and } h(x,y) = \frac{\gamma}{\rho}\left(\frac{1}{\gamma+\rho}\frac{x^{\gamma+\rho}-y^{\gamma+\rho}}{x^\gamma - y^\gamma} - \frac{1}{\gamma}\right). \tag{15}$$

From Lemma 1, one has when $\gamma > 0$ that

$$0 \leq \frac{1}{\gamma+\rho}\frac{x^{\gamma+\rho}-y^{\gamma+\rho}}{x^\gamma - y^\gamma} \leq \frac{1+\rho/\gamma}{\gamma+\rho} = \frac{1}{\gamma},$$

and when $\gamma < 0$,

$$\frac{1}{\gamma} \leq \frac{x^{\gamma+\rho}-y^{\gamma+\rho}}{x^\gamma - y^\gamma} \leq 0.$$

Replacing in (15) concludes the proof in the first case.
$(ii)$ Now, if $\gamma = 0$, one has

$$H_{0,\rho}(x) = \frac{1}{\rho}\left(\frac{x^\rho - 1}{\rho} - \ln(x)\right) \text{ and } h(x,y) = \frac{1}{\rho^2}\frac{x^\rho - y^\rho}{\ln(x/y)} - \frac{1}{\rho}.$$

The upper and lower bounds on $h$ can be then easily deduced from Lemma 1.
$(iii)$ Finally, when $\gamma = -\rho > 0$,

$$H_{-\rho,\rho}(x) = \frac{1}{\rho}\left(\ln(x) - \frac{x^\gamma - 1}{\gamma}\right) \text{ and } h(x,y) = -\frac{1}{\rho} - \frac{\ln(x/y)}{x^{-\rho} - y^{-\rho}}.$$

From Lemma 1,

$$0 \leq \frac{\ln(x/y)}{x^{-\rho} - y^{-\rho}} \leq -\frac{1}{\rho},$$

and the proof is completed. ■

**Lemma 3** *Under* **(C2)** *and (6),*

$$\lim_{x\to\infty} \sup_{(s,t)\in\tilde{\mathcal{D}}} |\overline{R}(s,t,x)| = 0.$$

*where $\tilde{\mathcal{D}} = \{(s,t) \in \mathbb{R}^2; s,t > 1 \text{ and } \neq t\}$.*

**Proof** – Remark first that

$$\overline{R}(s,t,x) = A^{-1}(x)\left[\frac{U(sx) - U(tx)}{(K_\gamma(s) - K_\gamma(t))a(x)} - 1\right] - h_{\gamma,\rho}(s,t)$$

and that
$$\frac{U(sx)-U(tx)}{a(x)} = \frac{a(tx)}{a(x)}\left[\frac{A(tx)}{A(x)}A(x)\left[R(s/t,tx)+H_{\gamma,\rho}(s/t)\right]+K_\gamma(s/t)\right].$$

From [23, Theorem 2.3.3] and [35, Proposition 0.5], one has

$$\lim_{x\to\infty}\sup_{t>1}\left|\frac{A(tx)}{A(x)}-t^\rho\right| = 0. \tag{16}$$

Furthermore, from [23, Corollary 2.3.5 and Theorem 2.3.6], one can take $a(x) = cx^\gamma(1-A(x)/\rho)^{-1}$ for some constant $c>0$. Hence, using (16) leads to

$$\frac{a(tx)}{a(x)} = t^\gamma\left(1+A(x)K_\rho(t)+o[A(x)]\right), \tag{17}$$

uniformly on $t>1$ as $x\to\infty$. Taking into account of (16) and (17), remarking that

$$t^{\gamma+\rho}H_{\gamma,\rho}(s/t) = (K_\gamma(s)-K_\gamma(t))(h_{\gamma,\rho}(s,t)-K_\rho(t))$$

and that $t^\rho\in(0,1)$ for all $t>1$ entail

$$\begin{aligned}\frac{U(sx)-U(tx)}{(K_\gamma(s)-K_\gamma(t))a(x)}-1 &= A(x)h_{\gamma,\rho}(s,t)+\mathcal{O}[A(x)]\frac{R(s/t,tx)}{K_\gamma(s/t)}\\ &+ o[A(x)]\frac{H_{\gamma,\rho}(s/t)}{K_\gamma(s/t)}+o[A(x)],\end{aligned}$$

uniformly on $s>t>1$. Since $h_{\gamma,\rho}$ is bounded from Lemma 2,

$$\lim_{x\to\infty}\sup_{s>t>1}\frac{|R(s/t,tx)|}{K_\gamma(s/t)} = 0$$

from (6) and since the function $s\mapsto H_{\gamma,\rho}(s)/K_\gamma(s)$ is bounded (the proof follows the same lines as the one of Lemma 2), one finally gets the expected result since $\overline{R}(s,t,x)$ is symmetric. ■

The following lemma is a key result for the proof of Propositions 4 and 5.

**Lemma 4** *Let $\Psi:(0,1)^2\to\mathbb{R}^+$ be a positive function such that for all $p\geq 0$ the integral,*

$$\mathcal{I} := \int_{(0,1)^2} K_p\left(\Psi(u,v)\right)dudv$$

*is absolutely convergent. For all $x>0$, let $\varphi_x:(0,1)^2\to[-1,\infty)$ a function satisfying*

$$\lim_{x\to\infty}\sup_{(u,v)\in\mathcal{D}}\varphi_x(u,v) = 0,$$

*where $\mathcal{D}=\{(u,v)\in(0,1)^2; u\neq v\}$. If, for all $x>0$, the integral*

$$\mathcal{J}_x := \int_{(0,1)^2} K_p\left(\Psi(u,v)(1+\varphi_x(u,v)\right)dudv$$

*is absolutely convergent then, as $x \to \infty$ and for all $p \geq 0$,*

$$K_p^{\leftarrow}(\mathcal{J}_x) = K_p^{\leftarrow}(\mathcal{I}) \left\{ 1 + \frac{1}{1+p\mathcal{I}} \int_{(0,1)^2} \Psi^p(u,v)\varphi_x(u,v)dudv\ (1+o(1)) \right\}.$$

**Proof** – Let $p \geq 0$. A first order expansion of the function $K_p$ leads to

$$K_p\left(\Psi(u,v)(1+\varphi_x(u,v)\right) = K_p\left(\Psi(u,v)\right) + \Psi^p(u,v)\varphi_x(u,v)(1+\zeta\varphi_x(u,v))^{p-1},$$

for some $\zeta \in (0,1)$. Since $\varphi_x(u,v)$ converges to 0 uniformly on $\mathcal{D}$ as $x \to \infty$, we obtain

$$\mathcal{J}_x = \mathcal{I} + \int_{(0,1)^2} \Psi^p(u,v)\varphi_x(u,v)dudv(1+o(1)). \tag{18}$$

For all $z > 0$, as $\varepsilon \to 0$, a first order Taylor expansion yields

$$\frac{K_p^{\leftarrow}(z+\varepsilon)}{K_p^{\leftarrow}(z)} = 1 + \frac{\varepsilon}{1+pz}(1+o(1)). \tag{19}$$

Taking $z = \mathcal{I}$ in (19) and

$$\varepsilon = \int_{(0,1)^2} \Psi^p(u,v)\varphi_x(u,v)dudv\ (1+o(1)) \to 0,$$

as $x \to \infty$, the conclusion follows from (18). ■

We finally recall a result on the asymptotic normality of $U$-statistics that is central in the proof of Theorem 1. The following lemma is a simplified version of [24, Theorem 7.1].

**Lemma 5** *Let $Y_1, \ldots, Y_m$ be independent identically distributed random variables and let $\Phi : \mathbb{R}^2 \to \mathbb{R}$ be a symmetric function. Define*

$$\mathcal{U}_m := \frac{1}{m(m-1)} \sum_{i \neq j} \Phi(Y_i, Y_j). \tag{20}$$

*If $\mathbb{E}[\Phi^2(Y_1,Y_2)] < \infty$, then, as $m \to \infty$, $m^{1/2}(\mathcal{U}_m - \mathbb{E}[\Phi(Y_1,Y_2)])$ converges to a centered Gaussian distribution with variance $4\text{Cov}[\Phi(Y_1,Y_2), \Phi(Y_1,Y_3)]$.*

From [24], the asymptotic variance is also equal to $4\mathbb{E}[\Psi^2(Y_1)]$ where $\Psi(x) := \mathbb{E}[\Phi(x,Y_2)] - \mathbb{E}[\Phi(Y_1,Y_2)]$.

## 7.2 Proofs of main results

**Proof of Proposition 1** – In the case where $p \geq 1$, the existence of the BC$_p$-TGV measure is a consequence of the triangular inequality. When, $p \in (0,1)$, the existence is deduced from the inequality $|u-v|^p \leq |u|^p + |v|^p$. Finally, in the situation where $p = 0$, the result is proved by remarking that the density of $|X - X^*|$ is bounded and thus $\mathbb{E}(|\ln|X - X^*||)$ exists as soon as $\mathbb{E}(|\ln|X||) < \infty$. ■

**Proof of Proposition 3** – For $p > 0$, the increasing property is a direct consequence of (1) and Holder's inequality. The case $p = 0$ is handled by a continuity argument. ■

**Proof of Proposition 4** − We start by remarking that for all $p \geq 0$,

$$\frac{G_X(\alpha;p)}{a(\alpha^{-1})} = K_p^{\leftarrow}\left(\int_{(0,1)^2} K_p\left(\Psi(u,v)[1+\varphi_{\alpha^{-1}}(u,v)]\right) dudv\right), \tag{21}$$

with $\Psi(u,v) := \left|K_\gamma(u^{-1}) - K_\gamma(v^{-1})\right|$ and

$$\varphi_{\alpha^{-1}}(u,v) := \frac{U(\alpha^{-1}u^{-1}) - U(\alpha^{-1}v^{-1})}{a(\alpha^{-1})(K_\gamma(u^{-1}) - K_\gamma(v^{-1}))} - 1. \tag{22}$$

Let us check that the functions $\Psi$ and $\varphi_{\alpha^{-1}}$ satisfy the assumptions of Lemma 4. First, by assumption, $G_X(\alpha;p)$ exists and thus for all $p \geq 0$,

$$\int_{(0,1)^2} \left|K_p\left(\Psi(u,v)[1+\varphi_{\alpha^{-1}}(u,v)]\right)\right| dudv < \infty.$$

Moreover, for all $p \geq 0$,

$$\int_{(0,1)^2} \left|K_p\left(\Psi(u,v)\right)\right| dudv < \infty,$$

since $p\gamma < 1$. Now, under **(C1)**,

$$\lim_{\alpha\to 0} \varphi_{\alpha^{-1}}(u,v) = \lim_{\alpha\to 0} \frac{\Delta(u^{-1},\alpha^{-1}) - \Delta(v^{-1},\alpha^{-1})}{K_\gamma(u^{-1}) - K_\gamma(v^{-1})} = 0.$$

Let us show that this convergence is uniform on $\mathcal{D} = \{(u,v) \in (0,1)^2; u \neq v\}$. Using the equality $K_\gamma(u^{-1}) - K_\gamma(v^{-1}) = v^{-\gamma}K_\gamma(v/u)$, one has

$$\begin{aligned}\Delta(u^{-1},\alpha^{-1}) - \Delta(v^{-1},\alpha^{-1}) &= \frac{a(v^{-1}\alpha^{-1})}{a(\alpha^{-1})}\left[\Delta\left(\frac{v}{u}, v^{-1}\alpha^{-1}\right) + K_\gamma\left(\frac{v}{u}\right)\right] \\ &- v^{-\gamma}K_\gamma\left(\frac{v}{u}\right).\end{aligned}$$

From condition (3), $a(v^{-1}\alpha^{-1})/a(\alpha^{-1}) = v^{-\gamma}(1+o(1))$ uniformly on $v > 1$, and thus,

$$\Delta(u^{-1},\alpha^{-1}) - \Delta(v^{-1},\alpha^{-1}) = v^{-\gamma}\Delta\left(\frac{v}{u}, v^{-1}\alpha^{-1}\right)(1+o(1)) + o\left(v^{-\gamma}K_\gamma\left(\frac{v}{u}\right)\right)$$

uniformly on $\mathcal{D}$. Hence,

$$\varphi_{\alpha^{-1}}(u,v) = \frac{\Delta(v/u, v^{-1}\alpha^{-1})}{K_\gamma(v/u)}(1+o(1)) + o(1),$$

which converges to 0 uniformly on $\mathcal{D}$ by the assumption (3). The conclusion follows by applying Lemma 4. ■

**Proof of Proposition 5** − We start with the expression of $G_X(\alpha;p)/a(\alpha^{-1})$ given in (21), proof of Proposition 4. Remark that, under **(C2)**,

$$\varphi_{\alpha^{-1}}(u,v) = A(\alpha^{-1})\left[h_{\gamma,\rho}\left(u^{-1},v^{-1}\right) + \overline{R}\left(u^{-1},v^{-1},\alpha^{-1}\right)\right].$$

From Lemma 3 and since the function $|h_{\gamma,\rho}|$ is bounded (see Lemma 2),

$$\varphi_{\alpha^{-1}}(u,v) = A(\alpha^{-1})h_{\gamma,\rho}\left(u^{-1},v^{-1}\right)(1+o(1)), \tag{23}$$

uniformly on $\mathcal{D}$. Remarking that

$$1 + p \int_{(0,1)^2} K_p\left(\Psi(u,v)\right) du dv = \theta^p(p,\gamma),$$

the conclusion follows by a direct application of Lemma 4. ■

**Proof of Theorem 1** − For all $p \geq 0$, let us consider the decomposition

$$\widehat{G}_{X,n}(\alpha_n;p) = a[U^{\leftarrow}(X_{n-k_n,n})]\widehat{\Theta}(\alpha_n;p), \tag{24}$$

where

$$\widehat{\Theta}(\alpha_n;p) := K_p^{\leftarrow}\left(\frac{1}{k_n(k_n-1)}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n} K_p\left(\frac{|X_{n-i+1,n} - X_{n-j+1,n}|}{a[U^{\leftarrow}(X_{n-k_n,n})]}\right)\right). \tag{25}$$

Let us introduce the standard exponential independent random variables $E_i := \ln U^{\leftarrow}(X_i)$, $i = 1,\ldots,n$ and focus on the first term of the right hand-side of (24). Denote by $D_n := E_{n-k_n,n} - \ln(n/k_n)$ to get

$$\frac{a[U^{\leftarrow}(X_{n-k_n,n})]}{a(n/k_n)} = \frac{a[(n/k_n)\exp(D_n)]}{a(n/k_n)}.$$

In view of $n/k_n \to \infty$, one has $k_n^{1/2} D_n \xrightarrow{d} \mathcal{N}(0,1)$, see for instance [23, Theorem 2.2.1]. From [23, Theorem 2.3.3], under **(C2)**, there exists $c_1 \in \mathbb{R}$ such that

$$\lim_{x\to\infty}\sup_{t\in\mathcal{K}}\left|A^{-1}(x)\left(\frac{a(tx)}{a(x)} - t^{\gamma}\right) - c_1 t^{\gamma} K_{\rho}(t)\right| = 0,$$

for all compact subset $\mathcal{K}$ of $(0,\infty)$. Letting $x = n/k_n$ and $t = \exp(D_n)$ in the previous expansion leads to

$$\begin{aligned}\frac{a[U^{\leftarrow}(X_{n-k_n,n})]}{a(n/k_n)} &= \exp(\gamma D_n)\left[1 + c_1 A(n/k_n) K_{\rho}(D_n)(1 + o_{\mathbb{P}}(1))\right] \\ &= \exp(\gamma D_n) + o_{\mathbb{P}}[A(n/k_n)].\end{aligned} \tag{26}$$

For $\gamma \neq 0$, the delta-method yields

$$k_n^{1/2}\left(\exp(\gamma D_n) - 1\right) \xrightarrow{d} \mathcal{N}(0,\gamma^2).$$

When $\gamma = 0$, we have shown that

$$k_n^{1/2}\left(\frac{a[U^{\leftarrow}(X_{n-k_n,n})]}{a(n/k_n)} - 1\right) \xrightarrow{\mathbb{P}} 0.$$

Let us now deal with the random variable (25). Recall that for all $i = 1,\ldots,k_n$, $X_{n-i+1,n} = U[\exp(E_{n-i+1,n})]$ and let $V_{i,n} := \exp[-(E_{n-i+1,n} - E_{n-k_n,n})] < 1$. Note that according to [23, Lemma 3.2.3], the random variables $\{V_{i,n},\ i = 1,\ldots,k_n\}$ are independent of $E_{n-k_n,n}$. Collecting (22) and (23) yields

$$\begin{aligned}\frac{U(\alpha^{-1}u^{-1}) - U(\alpha^{-1}v^{-1})}{a(\alpha^{-1})} &= \left(K_{\gamma}(u^{-1}) - K_{\gamma}(v^{-1})\right) \\ &\times \left[1 + A(\alpha^{-1}) h_{\gamma,\rho}(u^{-1},v^{-1})\left(1 + o(1)\right)\right],\end{aligned}$$

as $\alpha \to 0$ uniformly on $(u, v) \in \mathcal{D}$. Letting $u = V_{i,n}$, $v = V_{j,n}$ and $\alpha = \exp(-E_{n-k_n,n})$, a first order expansion of the function $K_p$ leads to

$$\begin{aligned}
& K_p\left(\frac{X_{n-i+1,n} - X_{n-j+1,n}}{a[U^{\leftarrow}(X_{n-k_n,n})]}\right) \\
= \;& K_p\left(\left|K_\gamma(V_{i,n}^{-1}) - K_\gamma(V_{j,n}^{-1})\right|\right) \\
+ \;& A[\exp(E_{n-k_n,n})]\left|K_\gamma(V_{i,n}^{-1}) - K_\gamma(V_{j,n}^{-1})\right|^p h_{\gamma,\rho}(V_{i,n}^{-1}, V_{j,n}^{-1}) \\
+ \;& o\left(A[\exp(E_{n-k_n,n})]\left|K_\gamma(V_{i,n}^{-1}) - K_\gamma(V_{j,n}^{-1})\right|^p\right),
\end{aligned}$$

for all $p \geq 0$ and uniformly on $(i, j) \in \{1, \ldots, k_n\}^2$ since the function $h_{\gamma,\rho}$ is bounded. Hence,

$$K_p\left(\widehat{\Theta}(\alpha_n; p)\right) = C_n(p) + A[\exp(E_{n-k_n,n})]B_n(p) + o_{\mathbb{P}}\left\{\tilde{C}_n(p)A[\exp(E_{n-k_n,n})]\right\},$$

with for all $p \geq 0$,

$$\begin{aligned}
C_n(p) &:= \frac{1}{k_n(k_n-1)}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n} K_p\left(\left|K_\gamma(V_{i,n}^{-1}) - K_\gamma(V_{j,n}^{-1})\right|\right), \\
\tilde{C}_n(p) &:= \frac{1}{k_n(k_n-1)}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n} \left|K_\gamma(V_{i,n}^{-1}) - K_\gamma(V_{j,n}^{-1})\right|^p, \\
\text{and} \quad B_n(p) &:= \frac{1}{k_n(k_n-1)}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n} \left|K_\gamma(V_{i,n}^{-1}) - K_\gamma(V_{j,n}^{-1})\right|^p h_{\gamma,\rho}(V_{i,n}^{-1}, V_{j,n}^{-1}).
\end{aligned}$$

The random variables $C_n(p)$, $\tilde{C}_n(p)$ and $B_n(p)$ are independent of $E_{n-k_n,n}$ and thus of $D_n$. Furthermore, let $F_1, \ldots, F_{k_n}$ (resp. $W_1, \ldots, W_{k_n}$) be standard exponential (resp. uniform) independent random variables. Rényi's representation theorem shows that

$$\begin{aligned}
\{V_{i,n},\ i = 1, \ldots, k_n\} &\stackrel{d}{=} \{\exp[-F_{k_n-i+1,k_n}],\ i = 1, \ldots, k_n\} \\
&\stackrel{d}{=} \{W_{i,k_n},\ i = 1, \ldots, k_n\}.
\end{aligned}$$

Since the sums are taken over all order statistics, for all $p \geq 0$,

$$\begin{aligned}
C_n(p) &:= \frac{1}{k_n(k_n-1)}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n} K_p\left(\left|K_\gamma(W_i^{-1}) - K_\gamma(W_j^{-1})\right|\right), \\
\tilde{C}_n(p) &:= \frac{1}{k_n(k_n-1)}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n} \left|K_\gamma(W_i^{-1}) - K_\gamma(W_j^{-1})\right|^p, \\
\text{and} \quad B_n(p) &:= \frac{1}{k_n(k_n-1)}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n} \left|K_\gamma(W_i^{-1}) - K_\gamma(W_j^{-1})\right|^p h_{\gamma,\rho}(W_i^{-1}, W_j^{-1}).
\end{aligned}$$

Hence, $B_n(p)$ is distributed as the $U$-statistic defined in (20) with $Y_i = W_i$ and $\Phi(y_1, y_2) = |K_\gamma(y_1^{-1}) - K_\gamma(y_2^{-1})|^p h_{\gamma,\rho}(y_1^{-1}, y_2^{-1})$ for all $p \geq 0$. Similarly, $C_n(p)$ (resp. $\tilde{C}_n(p)$) is distributed as the $U$-statistic defined in (20) with $Y_i = W_i$ and $\Phi(y_1, y_2) = K_p(|K_\gamma(y_1^{-1}) - K_\gamma(y_2^{-1}))$, (resp. $\Phi(y_1, y_2) = |K_\gamma(y_1^{-1}) - K_\gamma(y_2^{-1})|^p$). Besides, when $p\gamma < 1/2$, $\mathbb{E}(\Phi^2(W_1, W_2)) < \infty$ for all considered three functions, since

$|h_{\gamma,\rho}|$ is bounded (see Lemma 2), and therefore, one can apply Lemma 5 to establish that

$$\begin{array}{rcl} k_n^{1/2}\left(C_n(p) - K_p\left(\theta(p;\gamma)\right)\right) & \xrightarrow{d} & \mathcal{N}(0, 4\zeta_1^2(p;\gamma)), \\ k_n^{1/2}\left(\tilde{C}_n(p) - \theta^p(p;\gamma)\right) & \xrightarrow{d} & \mathcal{N}(0, 4\tilde{\zeta}_1^2(p;\gamma)), \end{array}$$

with, after straightforward calculations,

$$\begin{array}{rcl} \zeta_1^2(p;\gamma) & = & \displaystyle\int_0^1 \left(\int_0^1 K_p\left(\left|K_\gamma(u^{-1}) - K_\gamma(v^{-1})\right|\right) du\right)^2 dv - K_p^2(\theta(p;\gamma)), \\ \tilde{\zeta}_1^2(p;\gamma) & = & \displaystyle\int_0^1 \left(\int_0^1 \left|K_\gamma(u^{-1}) - K_\gamma(v^{-1})\right|^p du\right)^2 dv - \theta^{2p}(p;\gamma). \end{array}$$

Note that when $p = 0$, $\tilde{C}_n(0) = 1$ and $\tilde{\zeta}_1^2(0;\gamma) = 0$. Moreover, a sub-product of Lemma 5 is that $B_n(p) \xrightarrow{\mathbb{P}} \mathcal{B}(p;\gamma)$, for all $p \geq 0$, where $\mathcal{B}(p;\gamma)$ is defined in (5). Note also that, since $(k_n/n)\exp(E_{n-k_n,n}) \xrightarrow{\mathbb{P}} 1$ and since $A(\cdot)$ is regularly varying, $A[\exp(E_{n-k_n,n})] \overset{\mathbb{P}}{\sim} A(n/k_n)$. We finally get that, for $p \geq 0$,

$$k_n^{1/2}\left(K_p\left(\widehat{\Theta}(\alpha_n;p)\right) - K_p\left(\theta(p;\gamma)\right)\right) = \lambda\mathcal{B}(p;\gamma) + k_n^{1/2}\left(C_n(p) - \theta^p(p;\gamma)\right) + o_{\mathbb{P}}(1),$$

since $k_n^{1/2}A(n/k_n) \to \lambda \in \mathbb{R}$ by assumption. Using the delta-method, it follows

$$\begin{array}{rcl} k_n^{1/2}\left(\widehat{\Theta}(\alpha_n;p) - \theta(p;\gamma)\right) & = & \lambda\mathcal{B}(p;\gamma)\theta^{1-p}(p;\gamma) \\ & + & \theta^{1-p}(p;\gamma)k_n^{1/2}\left(C_n(p) - \theta^p(p;\gamma)\right) + o_{\mathbb{P}}(1). \end{array} \tag{27}$$

Collecting (26) and (27) and keeping in mind that $C_n(p)$, $\tilde{C}_n(p)$ and $D_n(p)$ are independent, expansion (24) yields

$$k_n^{1/2}\left(\frac{\widehat{G}_{X,n}(\alpha_n;p)}{a(n/k_n)} - \theta(p;\gamma)\right) \xrightarrow{d} \mathcal{N}\Big(\lambda\mathcal{B}(p;\gamma)\theta^{1-p}(p;\gamma); \; \theta^2(p;\gamma)\gamma^2 + 4\theta^2(p;\gamma)\zeta_1^2(p;\gamma)\zeta_2^2(p;\gamma)\Big), \tag{28}$$

for all $p \geq 0$, where $\zeta_2(p;\gamma) = (\theta(p;\gamma))^{-p}$. Recall that $\lfloor n\alpha_n \rfloor = k_n$ and thus that $n\alpha_n/k_n = 1 + \mathcal{O}(k_n^{-1})$. Since,

$$\frac{a(\alpha_n^{-1})}{a(n/k_n)} = 1 + \mathcal{O}(k_n^{-1}) + o[A(n/k_n)],$$

and Proposition 5 entails, for all $p \geq 0$,

$$\frac{G_X(\alpha_n;p)}{a(n/k_n)} = \theta(p;\gamma) + A(n/k_n)\mathcal{B}(p;\gamma)\theta^{1-p}(p;\gamma) + \mathcal{O}(k_n^{-1}) + o[A(n/k_n)]. \tag{29}$$

Combining (28) and (29) yields

$$k_n^{1/2}\left(\frac{\widehat{G}_{X,n}(\alpha_n;p) - G_X(\alpha_n;p)}{a(n/k_n)}\right) \xrightarrow{d} \mathcal{N}\left(0; \theta^2(p;\gamma)\gamma^2 + 4\theta^2(p;\gamma)\zeta_1^2(p;\gamma)\zeta_2^2(p;\gamma)\right)$$

and Proposition 4 concludes the proof:

$$k_n^{1/2}\left(\frac{\widehat{G}_{X,n}(\alpha_n;p)-G_X(\alpha_n;p)}{G_X(\alpha_n;p)}\right)\xrightarrow{d}\mathcal{N}\left(0;\gamma^2+4\mathcal{V}(p;\gamma)\right),$$

with $\mathcal{V}(p;\gamma):=\zeta_1^2(p;\gamma)\zeta_2^2(p;\gamma)$. ■

**Proof of Theorem 2** – As a first step, let us show that under (11),

$$k_n^{1/2}\left(\hat{\gamma}_n^*-\gamma,\frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})}-1\right)\xrightarrow{d}(\Lambda_1,\Lambda_2). \tag{30}$$

Let $(x,y)\in\mathbb{R}^2$ and introduce the Borel set $A_n:=\{p\hat{\gamma}_n<1\}$. One has

$$\begin{aligned}
&\mathbb{P}\left[\left\{k_n^{1/2}\left(\hat{\gamma}_n^*-\gamma\right)\le x\right\}\cap\left\{k_n^{1/2}\left(\frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})}-1\right)\le y\right\}\right]\\
=\ &\mathbb{P}\left[\left\{k_n^{1/2}\left(\hat{\gamma}_n^*-\gamma\right)\le x\right\}\cap\left\{k_n^{1/2}\left(\frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})}-1\right)\le y\right\}\cap A_n\right] &(31)\\
+\ &\mathbb{P}\left[\left\{k_n^{1/2}\left(\hat{\gamma}_n^*-\gamma\right)\le x\right\}\cap\left\{k_n^{1/2}\left(\frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})}-1\right)\le y\right\}\cap A_n^{\mathrm{C}}\right], &(32)
\end{aligned}$$

where $A_n^{\mathrm{C}}$ is the complement of $A_n$. From (11), $\hat{\gamma}_n$ converges in probability to $\gamma$. Hence, since $1/p-\gamma>0$, it follows that

$$\mathbb{P}(A_n^{\mathrm{C}})=\mathbb{P}\left(\hat{\gamma}_n-\gamma\ge\frac{1}{p}-\gamma\right)\to 0,$$

as $n\to\infty$. As a consequence, using the inequality,

$$\mathbb{P}\left[\left\{k_n^{1/2}\left(\hat{\gamma}_n^*-\gamma\right)\le x\right\}\cap\left\{k_n^{1/2}\left(\frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})}-1\right)\le y\right\}\cap A_n^{\mathrm{C}}\right]\le\mathbb{P}(A_n^{\mathrm{C}}),$$

we have shown that (32) converges to 0. Now, remark that under $A_n$, $\hat{\gamma}_n^*=\hat{\gamma}_n$. Hence,

$$\begin{aligned}
(31)\ =\ &\mathbb{P}\left[\left\{k_n^{1/2}\left(\hat{\gamma}_n-\gamma\right)\le x\right\}\cap\left\{k_n^{1/2}\left(\frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})}-1\right)\le y\right\}\right]\\
-\ &\mathbb{P}\left[\left\{k_n^{1/2}\left(\hat{\gamma}_n-\gamma\right)\le x\right\}\cap\left\{k_n^{1/2}\left(\frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})}-1\right)\le y\right\}\cap A_n^{\mathrm{C}}\right].
\end{aligned}$$

Under (11), the first term converges to $G(x,y)$ where $G$ is the cumulative distribution function of $(\Lambda_1,\Lambda_2)$ and the second term converges to 0. The proof of (30) is then complete.
The next step is based on the decomposition

$$\ln\frac{\widetilde{G}_{X,n}(\alpha_n;p)}{G(\alpha_n;p)}=\ln\frac{\theta(p;\hat{\gamma}_n^*)}{\theta(p;\gamma)}+\ln\frac{\hat{a}_n(\alpha_n^{-1})}{a(\alpha_n^{-1})}-\ln\frac{G(\alpha_n;p)}{\theta(p;\gamma)a(\alpha_n^{-1})}.$$

Since the function $\gamma\mapsto\theta(p;\gamma)$ is continuous for all $p\ge 0$, it follows that

$$\ln\frac{\theta(p;\hat{\gamma}_n^*)}{\theta(p;\gamma)}=\frac{1}{\theta(p;\gamma)}\left(\theta(p;\hat{\gamma}_n^*)-\theta(p;\gamma)\right)(1+o_{\mathbb{P}}(1)),$$

as $n\to\infty$. Now, $\gamma\mapsto\theta(p;\gamma)$ is continuously differentiable with $\dot{\theta}(p;\gamma)\neq 0$ for all $p\ge 0$ and $\gamma\in\mathbb{R}$

and therefore a first order Taylor expansion yields

$$\ln \frac{\theta(p;\hat\gamma_n^*)}{\theta(p;\gamma)} = \frac{\dot\theta(p;\gamma+\xi(\hat\gamma_n^*-\gamma))}{\theta(p;\gamma)}\,(\hat\gamma_n-\gamma)\,(1+o_{\mathbb{P}}(1)),$$

for some $\xi\in(0,1)$ as $n\to\infty$. The continuity of the function $\dot\theta(p;\cdot)$ and (30) lead to

$$k_n^{1/2}\ln\frac{\theta(p;\hat\gamma_n^*)}{\theta(p;\gamma)} \xrightarrow{d} \frac{\dot\theta(p;\gamma)}{\theta(p;\gamma)}\Lambda_1. \tag{33}$$

A direct application of (11) yields

$$k_n^{1/2}\ln\frac{\hat a_n(\alpha_n^{-1})}{a(\alpha_n^{-1})} = k_n^{1/2}\left(\frac{\hat a_n(\alpha_n^{-1})}{a(\alpha_n^{-1}}-1\right)(1+o_{\mathbb{P}}(1)) \xrightarrow{d} \Lambda_2. \tag{34}$$

Finally, Proposition 5 and a first order Taylor expansion of the function $\ln(1+\cdot)$ entail that

$$\ln\frac{G(\alpha_n;p)}{\theta(p;\gamma)a(\alpha_n^{-1})} = A(\alpha_n^{-1})\frac{\mathcal{B}(p;\gamma)}{\theta^p(p;\gamma)}(1+o(1)) = A(n/k_n)\frac{\mathcal{B}(p;\gamma)}{\theta^p(p;\gamma)}(1+o(1)), \tag{35}$$

since $A$ is regularly-varying. Collecting (33), (34) and (35) and remarking that

$$\ln\frac{\widetilde G_{X,n}(\alpha_n;p)}{G(\alpha_n;p)} = \left(\frac{\widetilde G_{X,n}(\alpha_n;p)}{G(\alpha_n;p)}-1\right)(1+o_{\mathbb{P}}(1))$$

prove the result. ■

**Proof of Corollary 1** − It is sufficient to show that the estimators $\hat\gamma_n^{(M)}$ and $\hat a_n^{(M)}(\alpha_n^{-1})$ satisfy condition (11) where $(\Lambda_1,\Lambda_2)$ is a Gaussian random vector with mean $\lambda(\mu_1(\gamma,\rho),\mu_2(\gamma,\rho))$ and covariance matrix

$$\begin{pmatrix} v_1(\gamma) & v_{1,2}(\gamma) \\ v_{1,2}(\gamma) & v_2(\gamma)\end{pmatrix}.$$

First, from [23, Lemma B.3.16], if $\gamma\neq\rho$, condition **(C2)** entails that there exists a function $B$ converging to 0 at infinity and such that

$$\lim_{x\to\infty} B^{-1}(x)\left(\frac{U(x)}{a(x)}\ln\frac{U(sx)}{U(x)} - K_{\gamma^-}(s)\right) = H_{\gamma^-,\rho'}(t),$$

with

$$\rho' = \begin{cases} \rho & \text{if } \gamma<\rho<0 \\ \gamma & \text{if } \rho<\gamma\le 0 \\ -\gamma & \text{if } (0<\gamma<-\rho \text{ and } \ell\neq 0) \\ \rho & \text{if } (0<\gamma<-\rho \text{ and } \ell=0) \text{ or } \gamma\ge-\rho>0.\end{cases}$$

Moreover, from [23, Lemma B.3.16], if $k_n^{1/2}A(n/k_n)\to\lambda\in\mathbb{R}$ then $k_n^{1/2}B(n/k_n)\to\tilde\lambda$ with

$$\tilde\lambda = \begin{cases} \lambda & \text{if } \gamma<\rho<0 \\ \left(\frac{\rho}{\gamma+\rho}\right)\lambda & \text{if } (0<\gamma<-\rho \text{ and } \ell=0) \text{ or } \gamma\ge-\rho>0 \\ 0 & \text{otherwise.}\end{cases}$$

One can thus apply [23, Corollary 4.2.2] to obtain that

$$k_n^{1/2}\left(\hat{\gamma}_n^{(M)}-\gamma, \frac{\hat{a}_n^{(M)}(\alpha_n^{-1})}{a(\alpha_n^{-1})}-1\right)$$

converges to a Gaussian distribution with desired mean and covariance. ■

**Proof of Theorem 3** − We start with the decomposition

$$\begin{aligned}
&\ln\left(\frac{\overline{G}^*_{X,n}(\beta_n;p)}{G_X(\beta_n;p)}\right)\\
=\ &\ln\left(\frac{\alpha_n}{\beta_n}\right)(\hat{\gamma}_n-\gamma)+\ln\left(\frac{\overline{G}_{X,n}(\alpha_n;p)}{G_X(\alpha_n;p)}\right)+\ln\left(\left(\frac{\alpha_n}{\beta_n}\right)^{\gamma}\frac{G_X(\alpha_n;p)}{G_X(\beta_n;p)}\right)\\
=\ &k_n^{-1/2}\ln\left(\frac{\alpha_n}{\beta_n}\right)\Lambda_{1,n}+k_n^{-1/2}\Lambda_{2,n}+\ln\left(\left(\frac{\alpha_n}{\beta_n}\right)^{\gamma}\frac{G_X(\alpha_n;p)}{G_X(\beta_n;p)}\right),
\end{aligned}$$

where $\Lambda_{1,n}\xrightarrow{d}\Lambda_1$ and $\Lambda_{2,n}\xrightarrow{d}\Lambda_2$. Let us focus on the third term of the right-hand side. From Proposition 5,

$$\frac{G_X(\alpha_n;p)}{G_X(\beta_n;p)}=\frac{a(\alpha_n^{-1})}{a(\beta_n^{-1})}\left\{\frac{1+A(\alpha_n^{-1})\mathcal{B}(p;\gamma)/\theta^p(p;\gamma)(1+o(1))}{1+A(\beta_n^{-1})\mathcal{B}(p;\gamma)/\theta^p(p;\gamma)(1+o(1))}\right\}$$

as $n\to\infty$. Since $A$ is regularly varying with index $\rho<0$ and since $\beta_n/\alpha_n\to 0$, one has that $A(\beta_n^{-1})=o(A(\alpha_n^{-1}))$ and consequently,

$$\frac{G_X(\alpha_n;p)}{G_X(\beta_n;p)}=\frac{a(\alpha_n^{-1})}{a(\beta_n^{-1})}\left\{1+\mathcal{O}(A(\alpha_n^{-1}))\right\}.$$

Moreover, from [23, Theorem 2.3.6], there exist two functions

$$a_0(x)=a(x)[1+\mathcal{O}(A(x))] \text{ and } A_0(x)=\frac{1}{\rho}A(x),$$

as $x\to\infty$ and such that for all $\varepsilon>0$ and $0<\delta<-\rho$,

$$\frac{1}{A_0(\alpha_n^{-1})}\left[\frac{a_0(\beta_n^{-1})}{a_0(\alpha_n^{-1})}-\left(\frac{\alpha_n}{\beta_n}\right)^{\gamma}\right]=\left(\frac{\alpha_n}{\beta_n}\right)^{\gamma}K_\rho\left(\frac{\alpha_n}{\beta_n}\right)+\mathcal{R}_n,$$

for $n$ large enough with $|\mathcal{R}_n|\le\varepsilon\,(\alpha_n/\beta_n)^{\gamma+\rho+\delta}$. As a consequence,

$$\begin{aligned}
&\left(\frac{\alpha_n}{\beta_n}\right)^{\gamma}\frac{G_X(\alpha_n;p)}{G_X(\beta_n;p)}\\
=\ &\left(\frac{\alpha_n}{\beta_n}\right)^{\gamma}\frac{a_0(\alpha_n^{-1})}{a_0(\beta_n^{-1})}\left\{1+\mathcal{O}(A(\alpha_n^{-1}))\right\}\\
=\ &\left\{1+A_0(\alpha_n^{-1})\left(K_\rho\left(\frac{\alpha_n}{\beta_n}\right)+\left(\frac{\alpha_n}{\beta_n}\right)^{-\gamma}\mathcal{R}_n\right)\right\}\left\{1+\mathcal{O}(A(\alpha_n^{-1}))\right\}.
\end{aligned}$$

Since $K_\rho(\alpha_n/\beta_n)\to-1/\rho$ and $(\alpha_n/\beta_n)^{-\gamma}|\mathcal{R}_n|\le\varepsilon\,(\alpha_n/\beta_n)^{\rho+\delta}\to 0$, it follows that

$$\left(\frac{\alpha_n}{\beta_n}\right)^{\gamma}\frac{G_X(\alpha_n;p)}{G_X(\beta_n;p)}=1+\mathcal{O}(A(\alpha_n^{-1}))=1+\mathcal{O}(A(n/k_n)),$$

in view of the regular variation property of $A$. Finally, $k_n^{-1/2}\ln(\alpha_n/\beta_n) \to 0$ yields

$$\ln\left(\frac{\overline{G}^*_{X,n}(\beta_n;p)}{G_X(\beta_n;p)}\right) = k_n^{-1/2}\ln\left(\frac{\alpha_n}{\beta_n}\right)\Lambda_{1,n} + k_n^{-1/2}\Lambda_{2,n} + \mathcal{O}(A(n/k_n)) \xrightarrow{\mathbb{P}} 0,$$

and the conclusion follows. ■


## Acknowledgements

This research was supported by the French National Research Agency under the grant ANR-19-CE40-0013-01/ExtremReg project. S. Girard gratefully acknowledges the support of the Chair Stress Test, Risk Management and Financial Steering, led by the French Ecole Polytechnique and its Foundation and sponsored by BNP Paribas, as well as the support of the French National Research Agency in the framework of the Investissements d'Avenir program (ANR-15-IDEX-02).


# References


[1] Albrecher, H., Beirlant, J., and Teugels, J. (2017). *Reinsurance: Actuarial and Statistical Aspects*, Wiley.

[2] Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion, *Journal of Banking and Finance*, **26**, 1505–1518.

[3] Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking and Finance*, **26**, 1487–1503.

[4] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, **9**, 203–228.

[5] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of extremes: Theory and applications*, Wiley.

[6] Bellini, F. and Di Bernardino, E. (2015). Risk Management with Expectiles, *The European Journal of Finance*. DOI:10.1080/1351847X.2015.1052150

[7] Box, G.E. and Cox, D.R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B*, **26**, 211–243.

[8] Breckling, J. and Chambers, R. (1988). M-quantiles, *Biometrika*, **75**, 761–772.

[9] Chen, Z. (1996). Conditional $L_p-$quantiles and their application to testing of symmetry in non-parametric regression. *Statistics and Probability Letters*, **29**, 107–115.

[10] Daouia, A., Girard, S. and Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles. *Journal of the Royal Statistical Society: Series B*, **80**, 263–292.

[11] Daouia, A., Girard, S. and Stupfler, G. (2019). Extreme M-quantiles as risk measures: From $L^1$ to $L^p$ optimization. *Bernoulli*, **25**, 264–309.

[12] Daouia, A., Girard, S. and Stupfler, G. (2020). Tail expectile process and risk assessment. *Bernoulli*, **26**, 531–556.

[13] Dekkers, A., Einmhal, J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *The Annals of Statistics*, **17**, 1833–1855.

[14] El Methni, J., Gardes, L. and Girard, S. (2018). Kernel estimation of extreme regression risk measures, *Electronic Journal of Statistics*, **12**, 359–398.

[15] El Methni, J. and Stupfler, G. (2017). Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, *Statistica Sinica*, **27**, 907–930.

[16] El Methni, J. and Stupfler, G. (2018). Improved estimators of extreme Wang distortion risk measures for very heavy-tailed distributions. *Econometrics and statistics*, 6, 129–148.

[17] Embrechts, P., Kluppelberg, C. and Mikosch, T. (2013). *Modelling extremal events for insurance and finance*, vol. 33. Springer Science & Business Media.

[18] Furman, E. and Landsman, Z. (2006). Tail variance premium with applications for elliptical portfolio of risks. *ASTIN Bulletin: The Journal of the IAA*, **36**, 433–462.

[19] Furman, E., Wang, R. and Zitikis, R. (2017). Gini-type measures of risk and variability: Gini shortfall, capital allocations, and heavy-tailed risks. *Journal of Banking and Finance*, **83**, 70–84.

[20] Gardes, L., Girard, S. and Stupfler, G. (2020). Beyond tail median and conditional tail expectation: extreme risk estimation using tail $L^p-$optimisation, *Scandinavian Journal of Statistics*, **47**, 922–949.

[21] Gómez-Déniz, E. and Calderín, E. (2014). Unconditional distributions obtained from conditional specifications models with applications in risk theory, *Scandinavian Actuarial Journal*, **7**, 602–619.

[22] Gómez-Déniz, E. and Calderín, E. (2015). Modeling insurance data with the Pareto arctan distribution, *ASTIN Bulletin*, **45(3)**, 639–660.

[23] de Haan, L., and Ferreira, A. (2006). *Extreme Value Theory: An introduction*, Springer Series in Operations Research and Financial Engineering, Springer.

[24] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution, *The Annals of Mathematical Statistics*, **19(3)**, 293–325.

[25] Holzmann, H. and Klar, B. (2016). Expectile asymptotics, *Electronic Journal of Statistics*, **10**, 2355–2371.

[26] Hua, L. and Joe, H. (2011). Second order regular variation and conditional tail expectation of multiple risks, *Insurance: Mathematics and Economics*, **49**, 537–546.

[27] Jones, M.C. (1994). Expectiles and M-quantiles are quantiles, *Statistics and Probability Letters*, **20**, 149–153.

[28] Konstantinides, D. (2018). *Risk Theory. A Heavy Tail Approach*, World Scientific Publishing.

[29] Leng, C. and Tong, X. (2014). Censored quantile regression via Box-Cox transformation under conditional independence, *Statistica Sinica*, **24**, 221–249.

[30] McNeil, A.J. (1988). Estimating the tails of loss severity distribution using extreme value theory, *ASTIN Bulletin*, **27(1)**, 117–137.

[31] McNeil, A.J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management*, Princeton University Press.

[32] Newey, W.K. and Powell, J.L. (1987). Asymmetric least squares estimation and testing. *Econometrica*, **55**, 819–847.

[33] Resnick, S. (2001). Discussion of the Danish data on large fire insurance losses, *ASTIN Bulletin*, **27(1)**, 139–151.

[34] Resnick, S. (2007). *Heavy-tail phenomena: probabilistic and statistical modeling.* Springer, New York.

[35] Resnick, S. (2008). *Extreme Values, Regular Variation, and Point Processes*, Springer.

[36] Rockafellar, R.T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions, *Journal of Banking and Finance*, **26**, 1443–1471.

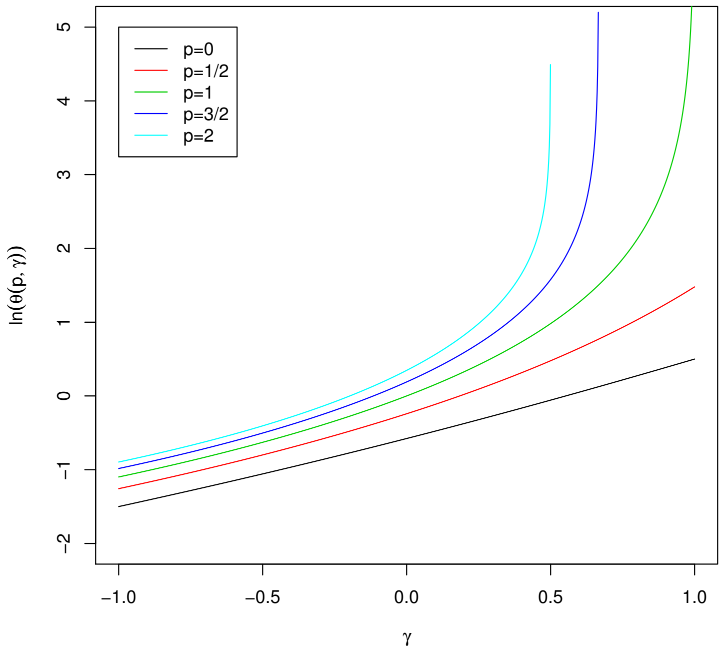
[37] Rockafellar, R. T., Uryasev, S. and Zabarankin, M. (2006). Generalized deviations in risk analysis. *Finance and Stochastics*, **10**, 51–74.

[38] Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. L. (2009). *Stochastic processes for insurance and finance*, vol. 505, Wiley.

[39] Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking and Finance*, **26**, 1519–1533.

[40] Valdez, E.A. (2005). Tail conditional variance for elliptically contoured distributions, *Belgian Actuarial Bulletin*, **5**, 26–36.

[41] Weissman, I. (1978). Estimation of parameters and large quantiles based on the $k$ largest observations, *Journal of the American Statistical Association*, **73**, 812–815.



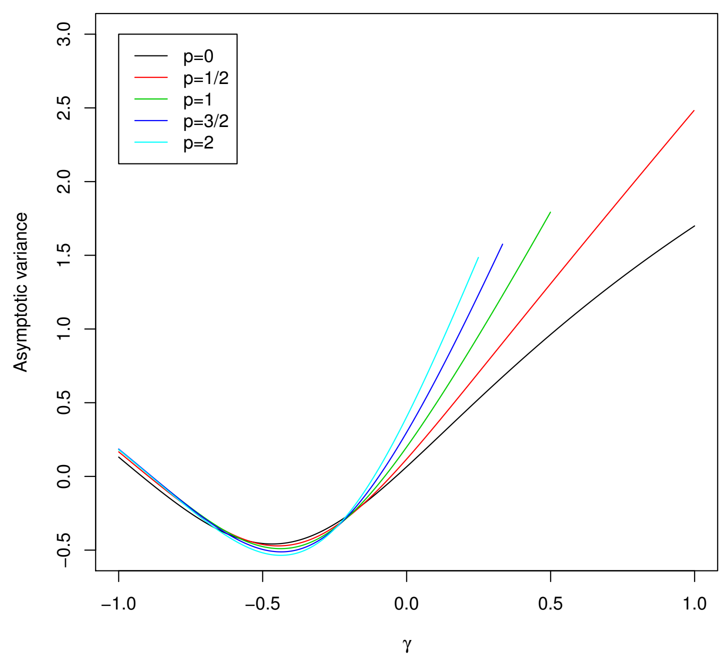




Figure 1: Graphes of $\gamma \mapsto \ln\theta(p;\gamma)$ under the constraint $p\gamma < 1$ (top panel) and $\gamma \mapsto \ln(\gamma^2 + 4\mathcal{V}(p;\gamma))$ under the constraint $p\gamma < 1/2$ (bottom panel) for $\gamma \in [-1, 1]$, $p \in \{0, 1/2, 1, 3/2, 2\}$. See Proposition 4 for the definition of $\theta(p;\gamma)$ and Theorem 1 for the definition of $\mathcal{V}(p;\gamma)$.

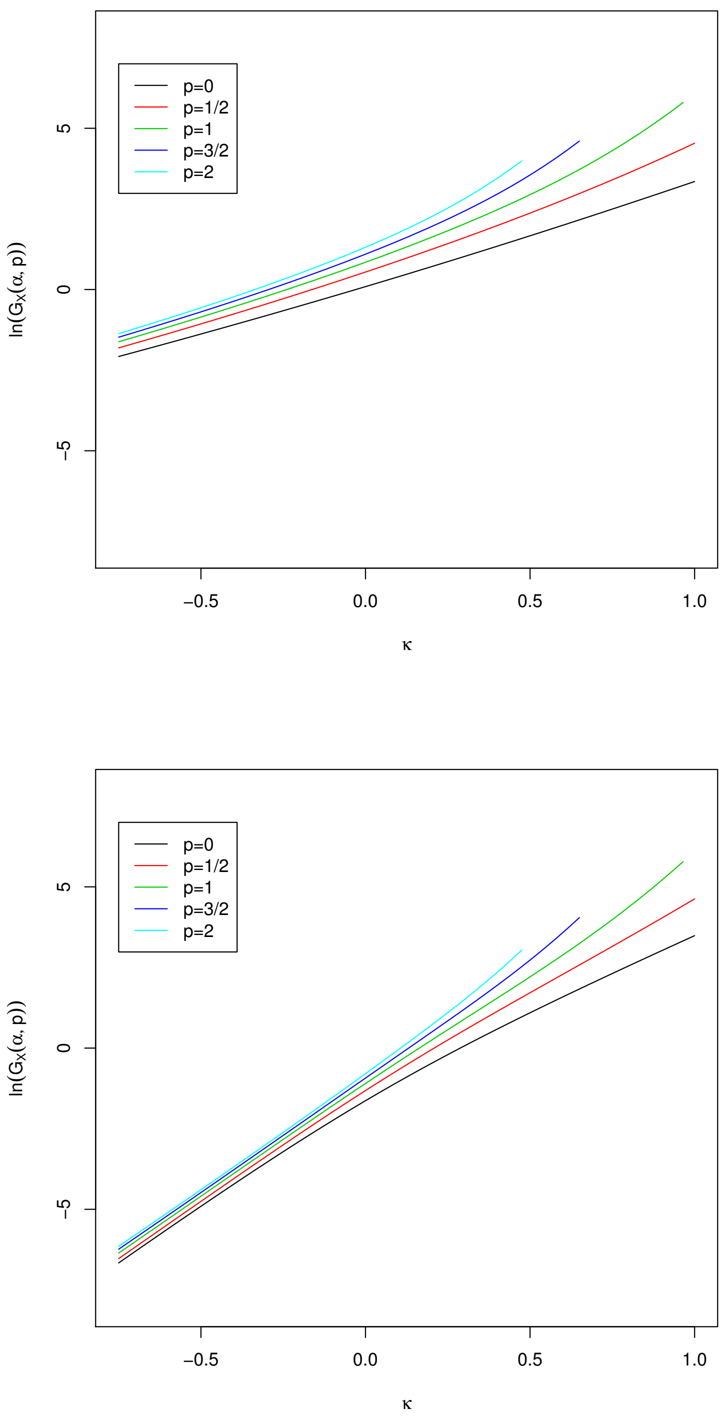



Figure 2: Graph of $\ln G_X(\alpha; p)$ as a function of $\kappa \in [-3/4, 1]$ for the distribution defined in (14) with $p \in \{0, 1/2, 1, 3/2, 2\}$, $\alpha = 0.04$, $c = 1/2$ (top panel) and $c = 2$ (bottom panel).

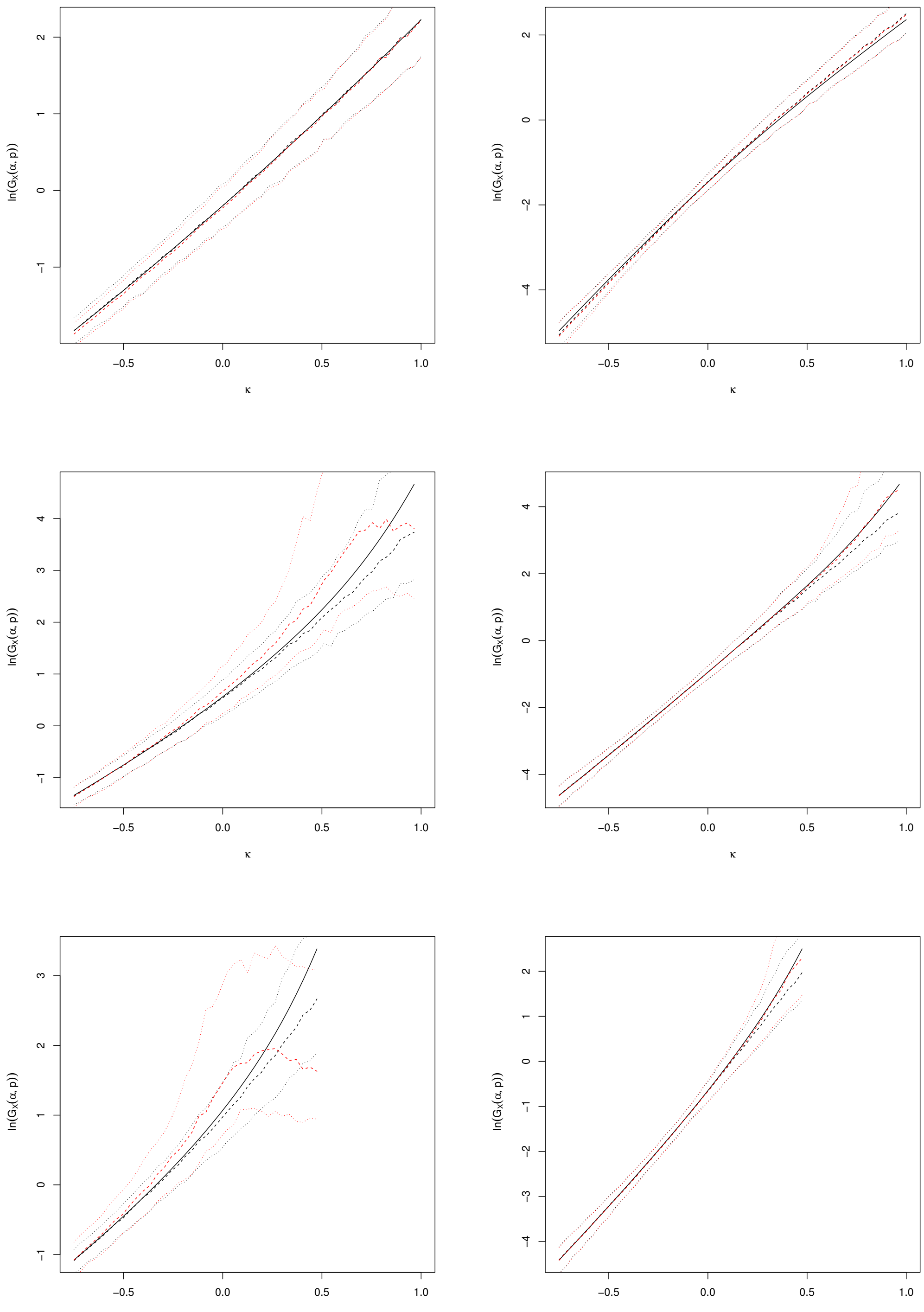



Figure 3: Simulated data, intermediate case with $n = 500$. The continuous line is the graph of the function $\kappa \mapsto \ln G_X(\alpha_n; p)$. The dashed lines represent the median of the logarithm of the $N$ realizations of the direct (black) and indirect (red) estimators. Dotted lines are the corresponding empirical quantiles of levels 5% and 95%. Top: $p = 0$, center: $p = 1$, bottom: $p = 2$. Left: $c = 1/2$, right: $c = 2$.

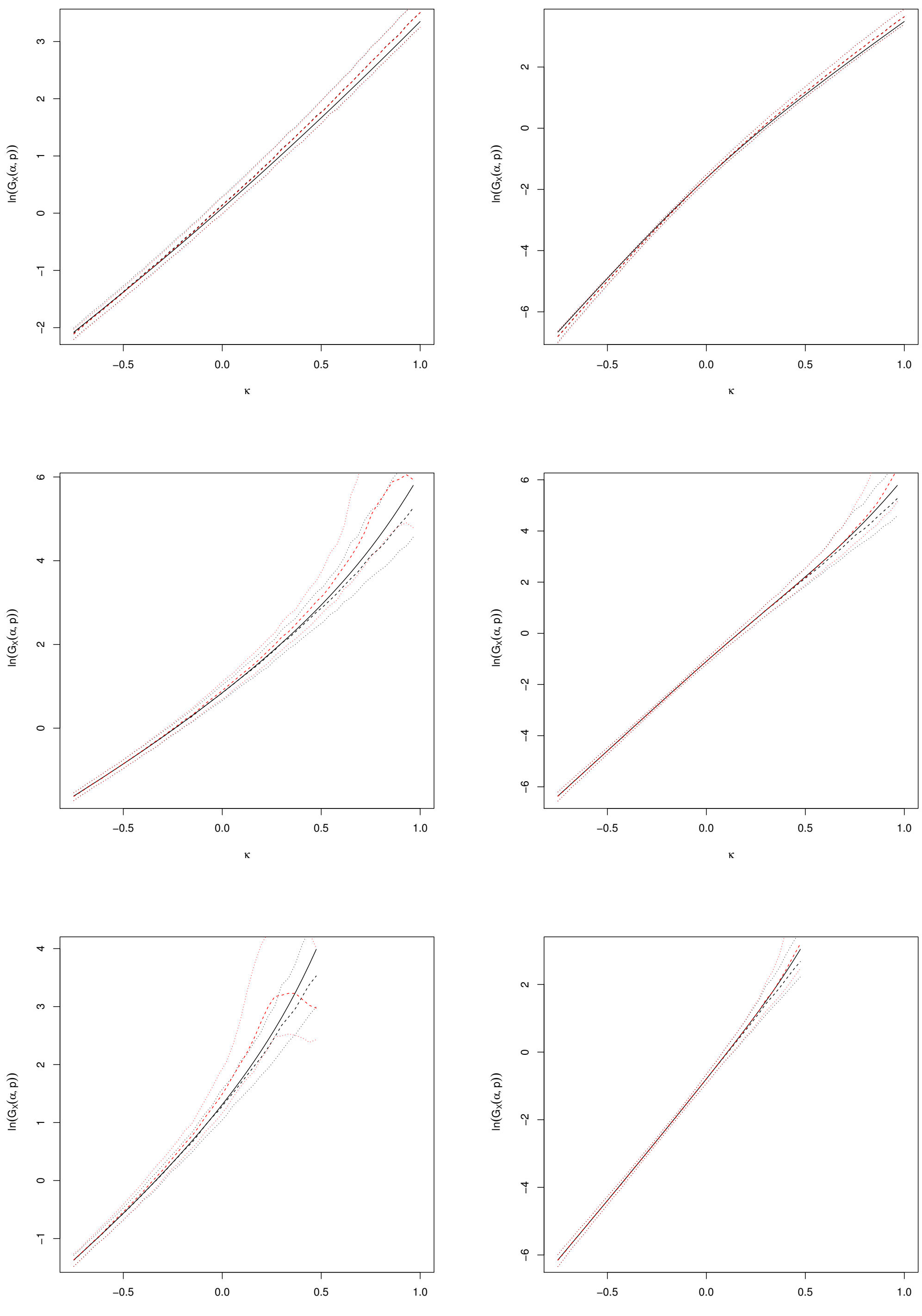



Figure 4: Simulated data, intermediate case with $n = 5000$. The continuous line is the graph of the function $\kappa \mapsto \ln G_X(\alpha_n; p)$. The dashed lines represent the median of the logarithm of the $N$ realizations of the direct (black) and indirect (red) estimators. Dotted lines are the corresponding empirical quantiles of levels 5% and 95%. Top: $p = 0$, center: $p = 1$, bottom: $p = 2$. Left: $c = 1/2$, right: $c = 2$.

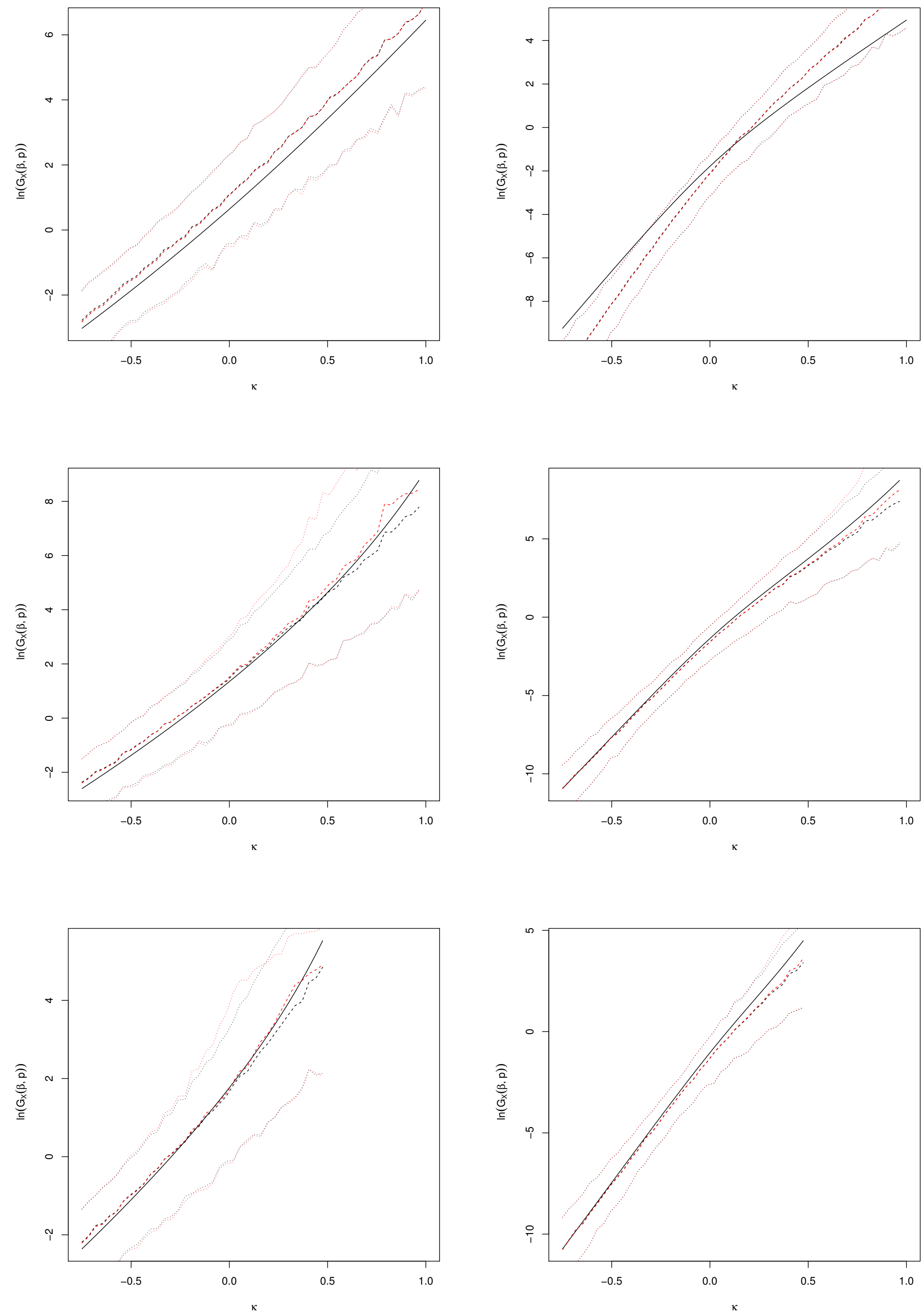



Figure 5: Simulated data, extreme case with $n = 500$. The continuous line is the graph of the function $\kappa \mapsto \ln G_X(\alpha_n; p)$. The dashed lines represent the median of the logarithm of the $N$ realizations of the direct (black) and indirect (red) extrapolated estimators. Dotted lines are the corresponding empirical quantiles of levels 5% and 95%. Top: $p = 0$, center: $p = 1$, bottom: $p = 2$. Left: $c = 1/2$, right: $c = 2$.

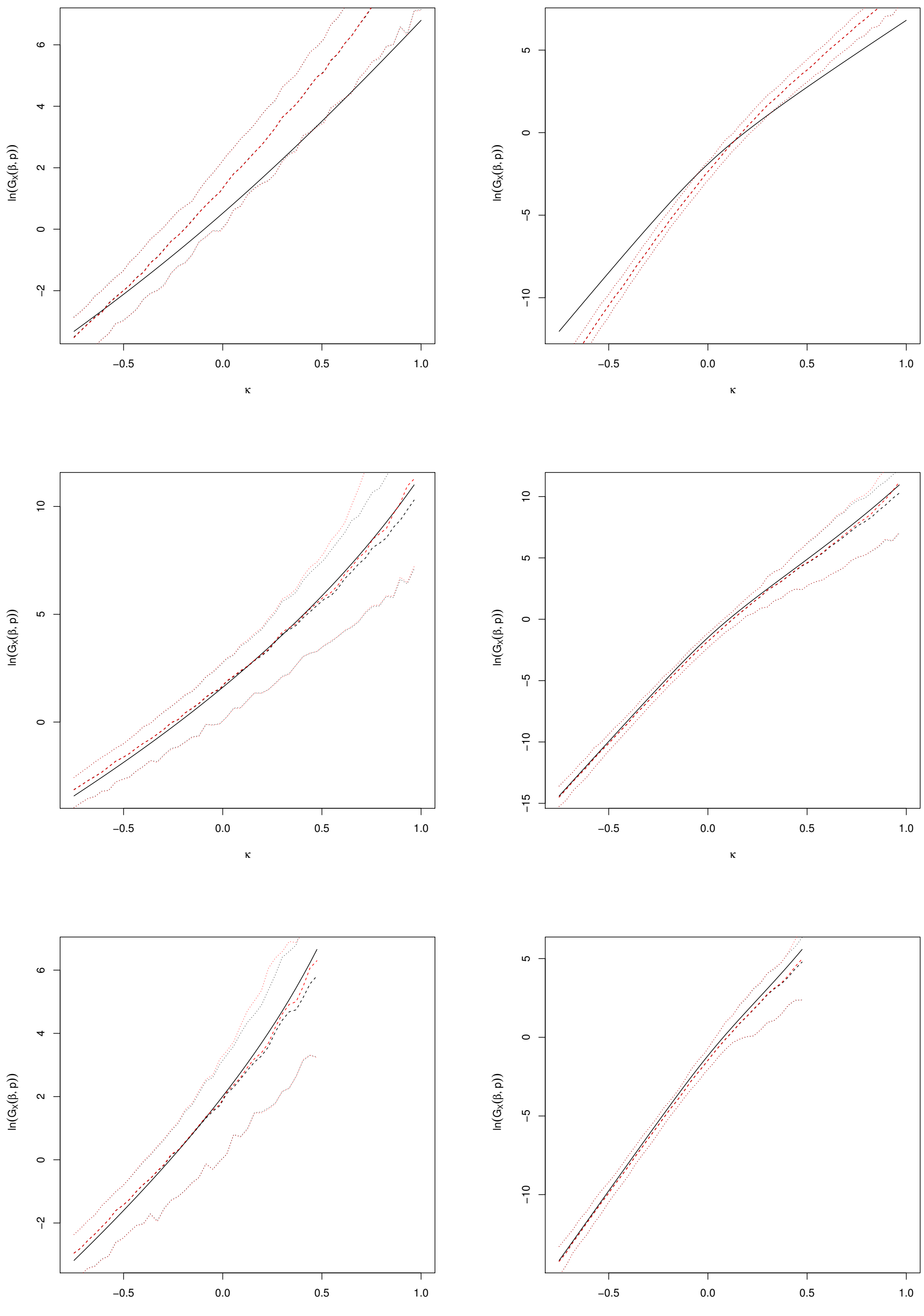



Figure 6: Simulated data, extreme case with $n = 5000$. The continuous line is the graph of the function $\kappa \mapsto \ln G_X(\alpha_n; p)$. The dashed lines represent the median of the logarithm of the $N$ realizations of the direct (black) and indirect (red) extrapolated estimators. Dotted lines are the corresponding empirical quantiles of levels 5% and 95%. Top: $p = 0$, center: $p = 1$, bottom: $p = 2$. Left: $c = 1/2$, right: $c = 2$.

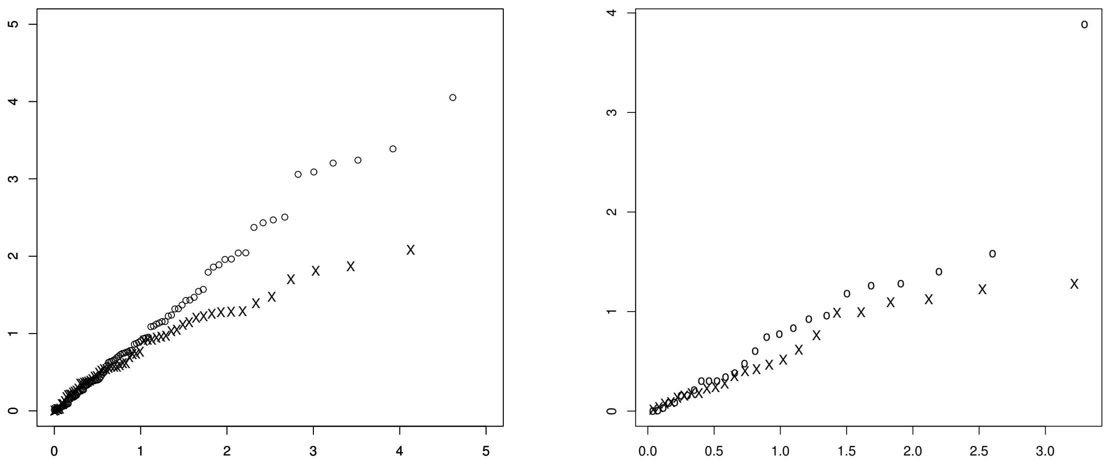


Figure 7: Left: Norwegian fire losses data set. QQ-plots obtained for years 1980 ($\times\times\times$) and 1985 ($\circ\circ\circ$). Right: Danish fire losses data set. QQ-plots obtained for years 1983 ($\times\times\times$) and 1980 ($\circ\circ\circ$).

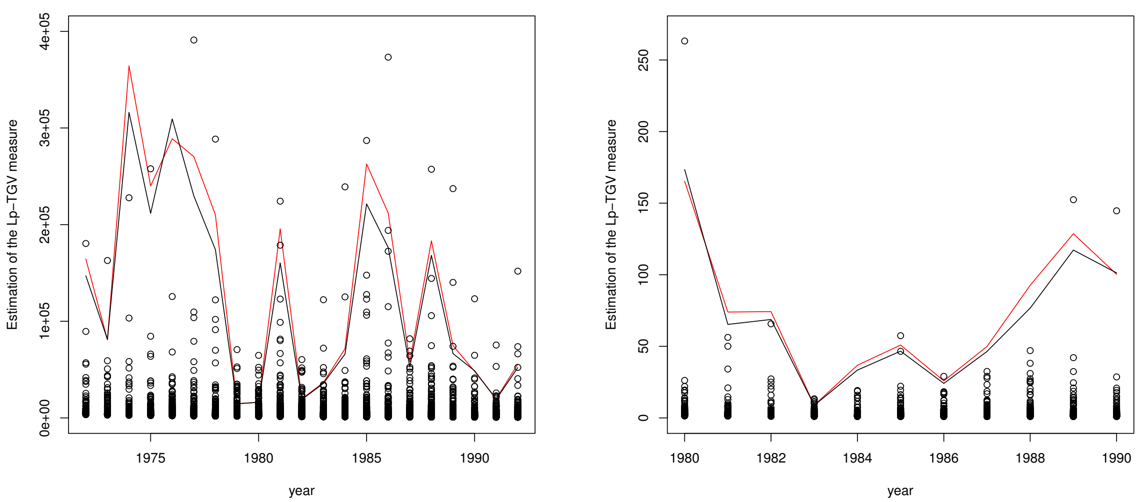



Figure 8: Estimation of the BC$_p$-TGV measure of level $\beta$ as a function of the year along with the values of the fire losses ($\circ\circ\circ$). In black, the direct extrapolated estimator; in red, the indirect extrapolated estimator. Continuous lines are used for the visualisation sake. Left: Norwegian fire losses data set with $\beta = 1/100$ and $p = \hat{p} \approx 0.694$. Right: Danish fire losses data set with $\beta = 1/150$ and $p = \hat{p} \approx 0.724$.